# A STUDY OF SINGLE- AND MULTI-LEVEL LOGISTIC REGRESSION MODELS USING REAL AND COMPUTER SIMULATED DATA

By

Mohamed Abdulla Kamali

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational
Psychology and Special Education

1992

# ABSTRACT

## A STUDY OF SINGLE- AND MULTI-LEVEL LOGISTIC REGRESSION MODELS USING REAL AND COMPUTER SIMULATED DATA

By

Mohamed Abdulla Kamali

This study provides a comparative analysis of the advantages and disadvantages associated with computer programs utilizing single- (i.e., SPSS) and multilevel logistic regression (MLR) estimation methods (i.e., VARCL, MULTILOGIT). Real and computer simulated data were employed in this study. Five different models of different complexity were investigated using real data. The simulated model included both a random intercept and random regression coefficient. The investigation considered random effects with both normal and t-distributions, various sample sizes of subjects within-groups, and different values of the random regression slope variance.

The findings drawn from running the SPSS, VARCL, and MULTILOGIT estimation programs using real data were: (1) The estimated regression coefficient for MULTILOGIT generally had a larger absolute value than both VARCL and SPSS. (2) The standard error estimates for both the within- and between-school variables regression coefficient for VARCL and MULTILOGIT were close and much larger than the SPSS estimates,

while the MULTILOGIT estimates were slightly larger than the VARCL estimates. (3) The estimate of variance-covariance components of the random effects for MULTILOGIT and VARCL were close. However, the MULTILOGIT estimates were generally larger absolute value than the VARCL estimates of the variance-covariance components. (4) There are several limitations of the MULTILOGIT program making its operation very restrictive.

The conclusions resulting from the SPSS, and VARCL estimation programs utilizing simulated data were: (1) Both the VARCL and SPSS estimates of $\gamma$'s were found to be significantly negatively biased and inconsistent. (2) The SPSS estimates of the standard error of macro parameters were significantly biased and inconsistent, while the VARCL estimates of the standard error of macro parameters were unbiased. (3) The probability of type I error rate under a true null hypothesis for the tests of the macro parameters $\gamma$'s were much smaller for VARCL than SPSS. However, both estimation method give unacceptable type I error rate (i.e., $p > .05$). (4) The VARCL estimates of $\tau_{00}$, $\tau_{11}$ and $\tau_{01}$ parameter were significantly negatively biased. However, the magnitude of the bias and MSE declined as the number of units within each group increased. (5) The VARCL estimates of the standard error for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{11}}$, and $\tau_{01}$ were significantly biased. However, the magnitudes of bias, and MSE were reduced as the sample size within each group increased.

**Dedicated to my father, mother, brothers, sisters,**

**and my wife and daughter**

**for their continuous love, and blessing.**

# ACKNOWLEDGEMENTS

I wish to thank Dr. Stephen W. Raudenbush, my academic advisor and chairperson of my doctoral dissertation for his guidance, insightful comments, and understanding. Working with him contributed greatly to my knowledge, development and understanding of applied statistics. Special thanks are also extended to Dr. Dennis Gilliland for his thoughtful and technical counseling, particularly during the final phases of the study. I also wish to thank the other members of my committee, Dr. William Schmidt and Dr. Habib Salehi.

I especially appreciate the support of the President and the Board of Trusteee of the United Arab Emirates University for sponsoring my studies and giving me this unique opportunity to obtain a doctoral degree.

Special thanks are given to my friend Ivan Filmer at Michigan State University for his help with editing this dissertation.

I wish to thank my father, mother, brothers, sisters and all the other members of my family for their love and support. Special thanks are also conveyed to my brother Adel Kamali for his help in computer programming during the summer of 1991.

Finally, I wish to express my deepest gratitude and thanks to my wife and daughter, Noor, for all their love, support, patience, and the many sacrifices they made so that this study could be completed.

TABLE OF CONTENTS

## LIST OF FIGURES

LIST OF TABLES

| Table | Page |
| --- | --- |

| Table | Page |
|---|---|

# CHAPTER I

# STATEMENT OF THE PROBLEM

## Introduction

One of the major difficulties in quantitative research in education is the departure from the normality of errors in statistical models. This departure may affect estimates, confidence intervals, and statistical conclusions. For example, researchers are frequently confronted with analyzing an observed behavior that is dichotomously scored, where '1' indicates an occurrence of the response, and '0' indicates the absence of a response. Models for such outcomes cannot have normal errors.

A second concern in educational research has been the appropriate analysis of multilevel data. For the past decade, there have been advances in multilevel data analysis with normally distributed outcomes. These have led to the development of several computer programs capable of analyzing data that have two or more levels of hierarchy. Some examples of such computer programs are GENMOD by Benjamin Hermalin based on Mason, Wong and Entwistle (1983); HLM by Bryk, Raudenbush, Seltzer and Congdom (1988) based on Raudenbush and Bryk (1986); ML3 by Rabash, Prosser and Goldstein based on Goldstein (1986); and VARCL by Longford based on Aitkin and Longford (1986) (all cited in Kreft and Kim, 1990). However, there have been concerns regarding the violations of the normality assumptions of the residuals in these programs. This is a concern especially for multilevel data that are dichotomously scored. Some examples of multilevel data structures having binary outcomes that are common in the field of education are:

1. Student repetition, where a value of "1" indicates that the student has repeated a grade, while a value of "0" indicates that the student has never repeated a grade;
2. Student persistence in school (dropouts vs. non-dropouts);
3. Student status of learning as mastery vs. non-mastery;
4. Student correct vs. incorrect response to a test item; and
5. Student attendance in college.

In all of these above-mentioned examples, there is a need to estimate the effect of both the student and school characteristics on student performance. Unfortunately, the logistic regression model, which was specifically designed for analyzing binary outcomes, is not capable of taking into account the inherent hierarchical structure of the data.

This has led to the advancement of several different approaches that take into account both (a) the binary response, and (b) multilevel data structure. However, the applications of these estimation methods in the field of educational research have been limited.

## Problem Statement

Several multilevel binary estimation methods have been proposed that take into account the multilevel data structure and binary outcomes. However, the advantages and the disadvantages of these proposed multilevel binary estimation methods have not been investigated.

Knowledge of the advantages and the disadvantages of these methods will help researchers make informed decisions about their applications in the field of education.

This study will identify the limitations of three different estimation methods that take into account both (a) the binary response and (b) multilevel data structure. In addition, the estimation method for the binary response of single level data will be investigated. The accuracy and statistical properties of the estimates for the multilevel logistic regression (MLR) model estimation methods and the single logistic regression model estimation procedure will be evaluated on real and computer simulated data.

## Purpose of the Study

The purpose of this study was to analyze and compare the single and multilevel logistic regression model estimation methods using four computer programs. The first three computer programs were based on different estimation methods designed for hierarchical data with binary outcomes, while the fourth program was based on single-level data with binary outcomes. The four computer programs used in this study were as follows:

1) Generalized Least Square method (Goldstein, 1990) using the ML3 program;
2) Quasi-likelihood method (Nelder and Pregibon, 1987) using the VARCL program;
3) Empirical Bayes estimation method (Wong and Mason, 1985) using the MULTILOGIT program; and
4) Maximum Likelihood estimation method using the SPSS program for single level data.

The following statistics were used to compare the four different estimation methods using data obtained from a national survey in Thailand:

1) the estimated regression coefficients for both student-level and school-level variables;
2) the estimated standard errors of the regression coefficients for both student-level and school-level variables; and
3) the estimated variance-covariance components of the random effects.

In addition, simulated data was generated to evaluate the parameter estimations obtained by using the four computer programs. The following criteria were used to gauge the accuracy of each estimation method:

1) the difference between the estimated regression coefficient and the true regression coefficient of

both within- and between-school variables;

2) the difference between the estimated variance component estimates and its true value; and

3) the difference between estimated standard error and the true standard error for: (a) within- and between-school regression coefficients, (b) the variance-covariance components of the random effects.

The above analyses were based on data generated according to three factors with two levels within each factor. These factors are the number of students within-school (small vs. large), the magnitude of the random slope variance (small vs. large), and distributions (both normal and t-) of the random effects terms. This resulted in a $2^2+2=6$ design matrix (the t-distributed random effects were investigated only under large random regression slope variance magnitude).

## The Need for the Study

Hopefully this study will not only identify the advantages, disadvantages, and estimation accuracy of existing multilevel binary estimation methods; but also inform researchers about the effect of student sample size within a school and the magnitude of the random regression slope. This could provide a researcher with the basis to decide which method is more appropriate in analyzing school-related data with various sample sizes when dealing with binary outcomes.

It is hoped that this study will narrow the gap between methodologists and practitioners in the field of multilevel binary data analysis.

It is further hoped that this study will also emphasize the need for researchers to concentrate on new areas of research in the development of multilevel binary data analysis, rather than duplicating an already developed approach with minimal changes.

## Research Question

The following research question guided the analysis of the data for this study:

Is there a difference in the accuracy of parameter estimation between i) the Generalized Least Squares, ii) the Quasi-likelihood, iii) the Empirical Bayes, and iv) the maximum likelihood (i.e., SPSS program) estimation methods, in relation to the multilevel logistic regression model? This research was applied to both real and simulated data. However, the accuracy of parameter estimates for the simulated data was evaluated for sixty groups according to the following conditions:

(i) a small sample of 10 students and a large sample of 60 students within a school,

(ii) the magnitude of .005% and 17.6% of the intercepts variance for the random regression slope, and

(iii) the normal distribution and the t-distribution of the random effects, $U_{oj}$ and $U_{1j}$.

## Multilevel Binary Models Used in the Study

In this study the multilevel logistic regression model was applied to both real and simulated data.

### The Real Data Models

The real data analysis was based on data obtained from a national survey in Thailand. This data was collected in 1988 by the National Education Commission of Thailand using a multistage cluster sample design. The entire sample consisted of 411 school principals, 3808 teachers, and 9768 sixth-grade students. Thus, the sample included both student- and school-level variables. For more information about the sample and the sample design, please refer to Raudenbush and Bhumirat (1989).

Because of MULTILOGIT computer program limitations, the real data analysis was based on 59 schools only. In order to compare the four computer programs, the same number of schools had to be used in each real data analysis.

Five different Multilevel Logistic Regression (MLR) models, from a simple to more complex models, were considered in real data analysis. These five models are shown in Chapter 3.

### The Simulated Model

The simulated analysis used the random intercept and random regression coefficient model. The simulated model contained one student-level variable and one school-level variable for the between-school model.

The within-group model was represented as

$$\alpha_{ij} = B_{oj} + B_{1j} X_{ij} \quad (1.1)$$

where

$\alpha_{ij}$ is the latent outcome variable which has been transformed to the log-odds by $\alpha_{ij}=\log(\theta_{ij}/(1-\theta_{ij}))$, $\theta_{ij}$ is the predicted probability of the subject (or student) i obtaining a value of 1 if he goes to the j-th group (or school), satisfying $P(Y_{ij}=1)=\theta_{ij}$ and $P(Y_{ij}=0)=1-\theta_{ij}$. This is assuming that $Y_{ij}$'s have a Bernoulli distribution with parameter $\theta_{ij}$ ( $E(Y_{ij})=\theta_{ij}$, $Y_{ij}|\theta_{ij}\sim$ Bernoulli($\theta_{ij}$));

$X_{ij}$ is the within-group level predictor for student i in the school j;

$B_{oj}$, $B_{1j}$ were random logistic regression coefficients across groups. $B_{1j}$ had both small and large magnitudes of variability between groups.

The between-group model was represented as

$$B_{oj} = \gamma_{oo} + \gamma_{o1} Z_j + U_{oj} \quad (1.2)$$

$$B_{1j} = \gamma_{1o} + \gamma_{11} Z_j + U_{1j} \quad (1.3)$$

where

$U_{kj}$ is the random effect where k=0 or 1, each with a mean of zero, and some variance $Var(U_{kj}) = \tau_{kk}$. For any pair of random effects $_k$ and $_{k'}$, $Cov(_{k},_{k'}) = \tau_{kk'}$;

$Z_j$ is the predictor for the school level;

$\gamma_{ko}$ is the overall intercept, and $\gamma_{k1}$ are the regression coefficients that capture the effects of school-level variables on the school regression coefficient $B_{kj}$. Thus the combined model was represented as

$$\alpha_{ij} = \gamma_{oo} + \gamma_{o1} Z_j + \gamma_{1o} X_{ij} + \gamma_{11} (Z_j * X_{ij}) + (U_{oj} + U_{1j} X_{ij}) \quad (1.4)$$

## Demonstrating the Model

The models were demonstrated using real and simulated data.

1. The Real Data: the analysis was based on 59 schools which were selected randomly from 411 schools. This data contained seven student variables (i.e. sex, dialect, SES, pre-primary education, repetition, having breakfast daily, having lunch daily) and five school level variables (i.e., urban/rural, central, north, south, mean SES). However, only one student variable (i.e., SES) was used.

This was due to the inability of the MULTILOGIT program to operate because of the small number of students within each school in the sample. The number of students within each school ranged between 8 and 37. This inability of the MULTILOGIT program is considered as one of its major weaknesses.

The sample used for the study did not have any missing data. This was to ensure that the data remained the same when analyzed using different computer programs.

2. The Simulated Data: An independent variable was produced having random regression slopes of .005% and 17.6% of the intercept's variance. In addition, the simulated data

contained data sets of 10 and 60 subjects (students) within a group (school) and, the normal and t-distribution of the random effects, $U_{oj}$ and $U_{1j}$. Finally, a single school-level variable ($Z_j$) was also produced.

This resulted in a simulated model having both a random intercept and a random regression slope.

## Research Procedures

Several techniques for analyzing multilevel binary data have been identified and a summary is shown in Table 2-1. Letters were sent to each of the researchers listed in Table 2-1 requesting their programs. The analyses in this study were based on three multilevel binary analysis programs (i.e., VARCL, ML3, MULTILOGIT programs) that were obtained. In addition, a single-level regression model estimation method (i.e., SPSS Program) was carried out.

## Summary

In spite of the development of several proposed methods of analyses for multilevel data with binary outcomes, the popularity of these programs and their applications in the field of educational research are limited. In addition, each of these programs has its own strengths and weaknesses. This study was aimed at conducting an analysis of several promising multilevel binary estimation methods consisting of the advantages of both the logistic regression model and the multilevel linear model. Often educational researchers are interested in analysis that takes into account the multilevel structure of the data and the nature of the binary responses of the students. The analysis of this study was based on comparing three multilevel and a single-level binary estimation method on real and simulated data. For the real data analysis, five different multilevel logistic regression (MLR) models (ranging from simple to more complex) were used. While the simulated model included both a random intercept and random regression coefficient, the generated (simulated) data considered the normal and t-distribution of the random effects, $U_{oj}$ and $U_{1j}$. In addition, the effect of small and large sample sizes of students within-school was investigated. Finally, a small and large magnitude of the random regression slope was also investigated in the simulated data.

# CHAPTER II

# REVIEW OF LITERATURE

This review of the literature will present an overview of the development of the logistic regression model, the multilevel linear model, and finally, the advancement of the multilevel logistic regression model.

## Logistic Regression Model

Concerns regarding the distribution of normal errors in the case of data with binary outcomes have led to the development of the logistic regression model. This model was specifically designed for analyzing binary data (Cox, 1970). For single level data with binary outcomes where ($Y_i$) takes the values "0" and "1" the expected value of $Y_i$ is

$$E(Y_i) = P(Y_i = 1) = \theta_i \qquad (2.1)$$

where

$\theta_i$ represents the probability of $Y_i$ equal to 1 (probability of success), and

$1-\theta_i$ represents the probability of $Y_i$ equal to 0 (probability of failure).

If the researcher wishes to investigate the dependence of $\theta_i$ on the independent variables ($X_1$ $X_2$ ... $X_p$), one possible way is to employ the ordinary linear regression technique where the model may be written as

$$Y_i = B_o + B_1X_{1i} + B_2X_{2i} + \ldots + B_pX_{pi} + \epsilon_i \tag{2.2}$$

where

$B_o$ represents the intercept; and

$B_1, B_2, \ldots, B_p$ represent the regression coefficients that characterize the relationship between the independent variables, $X_{1i}, X_{2i}, \ldots, X_{pi}$, and the dependent variable, $Y_i$.

The two basic assumptions of the linear regression model represented by Equation (2.2) are: (a) $\epsilon_i$ (error term) is a random variable with mean zero and variance $\sigma^2$, that is $E(\epsilon_i)=0$, $V(\epsilon_i)=\sigma^2$; and (b) $\epsilon_i$ and $\epsilon_j$ are not correlated, $i \neq j$ so that $Cov(\epsilon_i, \epsilon_j)=0$; thus the variance of $Y_i = \sigma^2$ and $Y_i$ and $Y_j$ where $i \neq j$ are not correlated. A further assumption which is not necessary for estimation, but is required in order to apply statistical tests such as the t- or F-tests, is that $\epsilon_i$ is a normally distributed random variable with mean zero and variance $\sigma^2$, that is $\epsilon_i \sim N(0, \sigma^2)$ (Draper and Smith, 1966). Thus $\epsilon_i$ and $\epsilon_j$ are not only uncorrelated but also independent.

However, the literature has cited several inadequacies and limitations of the linear regression model (Cox, 1970; Cox and Snell, 1989; Scheffe, 1959; Dunteman G., 1984; Hosmer and Lemeshow, 1989; Weisberg S., 1985; McCullgh and Nelder, 1989; Hanushek and Jackson, 1977; Clogg C., 1990; Efron, 1975; Anderson, 1980; Bull and Donner, 1987; Haberman, 1974, 1977). The main disadvantages of the linear regression model have been attributed to the violations of assumptions that $Y_i$'s are normally distributed with mean $\theta_i$ and variance $\sigma^2$ , and $\theta_i$ is linearly dependent on $X_i$'s. The limitations and the disadvantages of the above linear model could be summarized as follows:

1. It is quite possible that the predicted values of $\theta_i$ will exceed one or be a negative value.
2. Since $Y_i$ takes only the values 0 and 1, then $Y_i^2=Y_i$ and variance of $Y_i = \theta_i(1-\theta_i)$. This violates the assumption of the least squares estimate that variance $(Y_i) = \sigma^2$ (i.e., the assumption of homoscedasticity).

Using the least squares estimate could give us an unbiased estimate of $B_j$, but it is not an efficient estimator. This has led to the development of the logistic regression model which addressed the above problems by transforming the probability of success into a continuous variable that can take any value on the real line $(-\infty,\infty)$.

The logistic regression model is represented as follows:

$$\text{Logit}\ (\theta_i) = \text{Log}\ (\theta_i/(1-\theta_i)) = B_o + B_1X_{1i} + \ldots + B_pX_{pi} \qquad (2.3)$$

The logistic regression model is a sensible method for regression analysis of dichotomous data for two primary reasons. First, from a mathematical point of view, it is an extremely flexible and easily-used function. Second, it lends itself to a substantively meaningful interpretation (Hosmer and Lemeshow, 1989). It is the interpretation of the logistic regression coefficients that is the fundamental reason why logistic regression has proven to be such a powerful analytic tool for research (Breslow and Day, 1980; Alba R., 1987).

However, there are some disadvantages of the logistic regression model when dealing with data sets involving two or more levels of hierarchy. Assuming that the B's (regression slopes) are all fixed effects ignores the school (or group) effect on the variability between regression slopes. In fact, the concerns regarding usage of the single-level logistic regression model level are similar to the concerns about using the single-level regression model when analyzing data sets involving continuous outcomes and two or more levels of hierarchy.

## Multilevel Linear Model

There has been much educational research concerning the ability of a single-level regression model to deal with the hierarchical structure of data. In fact, most educational data can be seen as hierarchical where the lower level units are nested within the upper level units. For example, students are nested within classes, classes are nested within schools, schools are nested within districts, districts are nested within counties, and counties are nested within states.

Single-level analyses of data have led to several concerns regarding the unit of analysis and the violation of random sampling procedures (Langbein, 1977; Burstein, 1980; Kreft, 1987; Haney, 1980; Robinson, 1950; Alker, 1969; Hannan, 1971; Glass and Smith, 1979; Raudenbush and Bryk, 1988). This has led to the development of the multilevel linear model.

Within the field of educational research, this model not only illustrates the effect of student variables on the outcome but also the effect of school variables on both the aggregated student-dependent variable and the estimated within-school regression coefficients. This model may be represented by two equations.

For the within-school (i.e. group) model, we estimate a separate regression equation for each school:

$$Y_{ij} = B_{oj} + B_{1j}X_{1ij} + \ldots\ldots + B_{kj}X_{kij} + r_{ij} \qquad r_{ij} \sim N(0,\sigma^2) \qquad (2.4)$$

where

$i=1,2,\ldots,n_j$ students in school j,

$j=1,2,\ldots N$ schools, and

$k=1,2,\ldots,k$ independent variables within schools.

In this model $Y_{ij}$ is the response for student i in school j, $X_{kij}$ is the value of student-level independent variables k, and $r_{ij}$ is the random error. However, the assumption of $r_{ij} \sim N(0,\sigma^2)$ is violated due to the binary nature of the outcome variable. Coefficients $B_{1j}$ through $B_{kj}$ are regression coefficients that characterize the relationship within school j, and $B_{oj}$ is the intercept for each school.

The between-school (i.e. group) model is given by

$$B_{kj} = \gamma_{ko} + \gamma_{k1}Z_{1j} + \ldots + \gamma_{kp}Z_{pj} + U_{kj} \qquad U_{kj} \sim N(0,\tau) \qquad (2.5)$$

where

$U_{kj}$ the random effects $k=0,1,\ldots,k$ are assumed to be multivariate normal, each with a mean of zero, and some variance $Var(U_{kj}) = \tau_{kk}$. For any pair of random effect $_k$ and $_{k'}$, $Cov(_k,_{k'}) = \tau_{kk'}$. $Z_1,\ldots,Z_p$ are independent school variables, $\gamma_{ok}$

is the overall intercept, and $\gamma_{1K},\ldots,\gamma_{pK}$ are the regression coefficients that capture the effects of school-level variables on the school regression coefficient $B_{kj}$ adjusted for student intakes.

The key assumptions of the multilevel linear model are (a) the errors, $r_{ij}$, are normally distributed; and (b) within-group regression coefficients (B's) are assumed to be multivariate normally distributed. Both will be violated if the outcome in the within-group model is dichotomous (Leonard T., 1972b).

## Multilevel Logistic Regression Model

In the case of binary response data, there have been concerns regarding the violations of normality assumptions of the residuals. These concerns have been indicated by several researchers (e.g. Mason et al. 1984; Clogg et al. 1990; Leonard, 1972a, 1972b, 1975; Anderson and Aitkin, 1985; Stiratelli et al. 1984; Wong and Mason, 1985; Raudenbush and Bryk, 1986; Raudenbush, 1988; Goldstein, 1987; Braun, 1989; Lindely and Smith, 1972 in discussion p. 24). Recently, Longford (1990) has expressed this concern by stating:

> Normal distribution of the random terms in multilevel analysis is an important restrictive assumption. Much of the observational data in the social sciences are inherently discrete, and in the extreme, binary (e.g., Yes/No responses to survey questions). For such data the normal linear multilevel analysis is not appropriate not only because of the violation of the assumption of normality, but also because we usually wish to use a

> nonlinear scale such as the logit for binomial data, logarithm for Poisson data, etc. It is therefore desirable to have an extension of the multilevel methods for a wider class of distributional assumptions, which would at the same time be an extension of the methods for regression analysis of independent non-normally distributed data. (p. 2)

Mason et al, (1984) have also indicated a similar concern for estimation methods that account for discretion:

> The methodology presented in this chapter by no means exhausts the subject of multilevel estimation. There is a need for estimation procedures to handle discrete micro response variables. (p. 100)

After comparing four major computer packages for multilevel linear regression techniques (i.e. GENMOD, HLM, ML2 and VARCL), Kreft et al. (1990a, 1990b) found that the assumption of linearity in existing techniques and the assumption of normality of residuals were the limitations of some existing multilevel techniques.

In fact, the inadequacy of the multilevel linear model due to the violation of the normality assumption of the residuals could contribute to the following concerns:

1. Inadequacy in estimates of the within-school (or group) model variance, $\sigma^2$. Since $Y_{ij}$ takes only the values 0 and 1, then $Y_{ij}^2=Y_{ij}$ and variance of $Y_{ij}$ = $\theta_{ij}(1 - \theta_{ij})$. This violates the assumption of the estimate that variance$(Y_i)=\sigma^2$ could effect the

estimated standard error of the within-school coefficient (B's) when the hypothesis, $H_o$: B=0 is tested. The above concerns have also been stated by Raudenbush and Bryk (1986):

> There has been little empirical work on the consequences of violating normal distribution assumptions in HLM, but we suspect that problems are most likely to occur in estimates of the model variances, $\sigma^2$ and $\tau$, and in hypothesis-testing application. (p. 14)

2. It is possible that since $Y_{ij}$ takes values 0 and 1, the obtained fitted values of the regression parameter for the linear regression model would not satisfy the condition that $0 \leq E(Y|X,Z) = \theta_{ij} \leq 1$.

The research literature has also shown several different approaches to overcome the above concerns and to take into account both the binary response and multilevel data structure. (These approaches are summarized in Table 2-1). The majority of these approaches are based on the idea that new techniques should consist of the advantages of both the logistic regression model and the multilevel linear model.

## Development of Multilevel Logistic Regression Model

Initial concerns regarding using multilevel linear model analysis (Equations 2.4 and 2.5) in the case of binary outcomes were indicated by Leonard (cited in a commentary by Lindley and Smith (1972), where he stated:

> I would like to make a few remarks about the possible extension of the excellent ideas expressed in this paper to

situations where the exchangeable parameters cannot be considered to be normally distributed. In such circumstances, a good procedure is usually to transform the parameters in such a way that the normality assumption is more realistic for the new parameters. (p. 24)

Table 2-1.--Several multilevel binary analysis techniques.

| Author | Methods of estimation | Methodological Reference |
|---|---|---|
| Leonard T. | Bayesian* | Lenoard (1972a, 1972, 1975) |
| Chamberlain G. | Maximum Likelihood | Chamberlain (1980) |
| Wong G. & Mason N. | Empirical Bayes | Wong & Mason (1985) |
| Stiratelli et al. | Maximum Likelihood for fixed effect & variance component. Empirical Bayes estimate of random effect. | Stiratelli et al. (1984) |
| Anderson D.& Aitkin M. | Maximum Likelihood | Anderson & Aitkin (1985) |
| Clogg C. et al. | Bayesian* | Clogg C. et al. (1990) |
| Longford N. | Quasi-likelihood | Nelder & Pregibon (1987) |
| Goldstein H. | Generalized Least Square | Goldstein H. (1990) |
| Korn E. & Whittmore A. | Maximum Likelihood* | Korn & Whittmore (1979) |

* The goal of these methods is to combine the regression coefficients across groups into single coefficients for each of the covariates (i.e.borrowing strength).

Leonard suggested the use of Log-odds transformation (where $E(Y_i) = \theta_i$ for i=1,2,..,n, $Y_i$ being independent and binomial distributed with parameter $\theta_i$) where $\alpha_i = \log(\theta_i/(1-\theta_i))$, with the assumption that $\alpha_i \sim N(\mu,\sigma^2)$, where $\mu$ is uniformly distributed and $\sigma^2$ possesses an inverse $\chi^2$ distribution when $\sigma^2$ is known. The main point for this transformation was to estimate $\theta_{i.}$ adjusted for each group and the overall mean.

Using the Bayesian estimation procedure, Leonard (1972a, 1972b, 1975) extended his ideas for binary data with an application to the prediction of college (i.e. group) success rates ($Y_{ij}$). In this case, student college grades corresponded to the pass/fail situation with $i=1,2,...,n_j$ number of student within-college (or group), and $j=1,..,N$ colleges. By combining the available information ($X_{kij}$'s, student independent variables, student test scores on k different scales previous to college entry) from all the colleges to obtain predictors, more reliable results were produced than if the predictors were based only on information from one college.

Leonard assumed that $Y_{ij}$'s are mutually independent and have a Bernoulli distribution with parameter $\theta_{ij}$ ($E(Y_{ij})=\theta_{ij}$, $Y_{ij}|\theta_{ij}\sim$ Bernoulli($\theta_{ij}$)). For the within-college model, a separate logistic regression equation was estimated for each college (the symbols have been modified in order to be consistent with previously used symbols)

$$\alpha_{ij} = B_{oj} + B_{1j}X_{1ij} + \ldots\ldots + B_{kj}X_{kij} \qquad (2.6)$$

where

$\alpha_{ij}$ is the latent outcome variable which has been transformed to the log-odds by $\alpha_{ij}=\log(\theta_{ij}/(1-\theta_{ij}))$;

$\theta_{ij}$ is the predicted probability of the student i obtaining a degree if he goes to the j-th college, satisfying $P(Y_{ij}=1)=\theta_{ij}$ and $P(Y_{ij}=0)=1-\theta_{ij}$, given the $Y_{ij}$, binary outcome for student i (i.e., pass/fail) with college j;

$B_{oj}$ through $B_{kj}$ are within-college level logistic regression coefficients; and

$X_{kij}$ is the within-college level predictor k for student i in the school j. This is assuming that the within-college logistic regression coefficient $B_j$'s are exchangeable.

In addition, Leonard made two assumptions regarding the prior distribution for the vectors of the logistic regression coefficients (i.e. $B_{oj},...,B_{kj}$).

(a) Given $\mu_B$ and $H_B$, the $(B_{oj},...,B_{kj})$ are independent and have multivariate normal distributions with common mean vector $\mu_B$ and precision matrix $H_B$.

(b) the mean vector $\mu_B$ is uniformly distributed over (K) dimensional real space. Also $WH_B$ is independent of $\mu_B$ and has a Wishart distribution with W degrees of freedom and parametric matrix $Z_B^{-1}$.

Leonard applied a Bayesian approach (estimating the joint posterior modes for $B_{oj},...,B_{kj}$) with Newton's iterative procedure in order to obtain within-college coefficients. However, he encountered a problem in finding a starting value for the within-college coefficients (the $B_j$'s). In addition, his model did not include school (or college) level variables at the second stage of the model. But the aim of the model was to combine the available information (i.e. within colleges) in all colleges to obtain predictors which were more reliable than if the predictors were based only on the information from one college (i.e. borrowing strength).

However, several researchers have indicated a concern regarding the use of the approximate normal distribution for the posterior distribution (Laird, 1978; Laird and Louis, 1982; Geisser, 1984). As Laird and Louis (1982) indicated:

> The normal approximation has been used (see Leonard, 1975; Laird, 1977), but no indication of its validity was given in these papers. For the censored exponential, the normal approximation fails to account for the skewness of the gamma; for the 99.9 percent confidence interval it produces a negative left endpoint. (p. 199)

Chamberlain (1980) studied a random effects model for binary outcomes in which the intercepts were assumed to follow a distribution (i.e, random intercept for the within-group model), while other logistic regression coefficients were fixed across groups (in his study the groups were the individuals) in order to capture group differences.

Thus, for the within-group model (similar to Equation 2.6, with assumed fixed regression slope's, $B_1, \ldots, B_k$).

$$\alpha_{ij} = B_{oj} + B_1 X_{1ij} + \ldots\ldots + B_k X_{kij} \qquad (2.7)$$

While the between group model is

$$B_{oj} = \gamma_{oo} + \gamma_{o1} X_{1j} + \ldots\ldots + \gamma_{op} X_{pj} + U_{oj} \qquad (2.8)$$

(Assuming that $U_{oj}$ are independent and identically distributed.)

In the above analysis, Chamberlain's concern was the within-group estimator $(B_1,...,B_k)$. Thus he used a random intercept (i.e., $B_{oj}$) model in order to capture omitted variables that were group specific. Maximum likelihood procedures were used to estimate the model's parameters.

Wong and Mason (1985) introduced a multilevel binary model called a "Hierarchical Logistic Regression Model" which combined the advantages of the multilevel linear model and the logistic regression model when dealing with binary outcomes. They used the logistic regression model as the within-group model (i.e. within-school) and the multilevel linear model (Equation 2.5) for the between-group model. Thus, the within-group model is similar to Leonard's (1972b, 1975) model (Equation 2.6), where a separate logistic regression coefficient was estimated for each group. The between-group model (Equation 2.5) represented the effect of group variables on the estimated logistic regression coefficients for each group. This allowed the specification of the effect of the upper level (i.e., group membership) on the lower level of the hierarchy.

In fact, the major difference between Equation 2.3 (the fixed effect model) and the above two-stages model of Wong and Mason (also known as the mixed model) is the presence of the error terms in the between-group model (i.e., $U_{oj},U_{1j},...,U_{kj}$). Thus, if the between-group error terms are suppressed, the multilevel logistic regression model becomes a logistic regression model.

In deriving their mixed model, Wong and Mason have made two main assumptions: (a) within-group regression coefficients ($B_j$'s) are assumed to be normally distributed over group membership; (b) there is flat prior in the between-group coefficients given by $\gamma \sim N(m,\Sigma)$, $\Sigma^{-1}$ --->0. In addition, it is assumed "...that the $n_j$ are large enough to permit estimation of all $B_j$." (p. 514). In fact, the above hierarchical logistic regression model could be viewed as a classical discrete mixed model with fixed effects, $\gamma$, and random effects, $U_{kj}$. Empirical Bayes estimation procedures were used to estimate the parameters of the model where $\tau$ was estimated by the indirect Maximum Likelihood estimator using the EM alogarithm (Dempster et al. 1977, 1981). This was because of the difficulties in direct numerical maximization of the likelihood. Approximate posterior interval estimates were used to estimate $\gamma$'s and $B_j$'s.

In spite of the advantages of the proposed Wong and Mason model which takes into account the multilevel structure of the data and the nature of the binary student response, there are several concerns. These concerns are summarized as follows:

(a) The above model requires a large sample size within each group in order to permit the estimation of all $B_j$'s (Wong and Mason, 1985 p.514). This is often not the case in the field of education where the number of students within each school is small. In their study, Wong and Mason used the countries as the unit of the analysis in the second stage of their model (i.e. between-group model);

(b) Another concern that is also indicated by Wong and Mason is that, "Extensive exploration of the data using the computational procedure described here is costly for large data sets, because of the slow convergence of the EM algorithm for variance and covariance component problems." (p. 522)

(c) Raudenbush (1988) also indicated some concerns regarding the estimation procedures of Wong and Mason (1985):

> ...the data are binomial distributed conditional on the logistic regression coefficients for each country. These "random coefficients" are then assumed normal. Since the normal is not the conjugate prior for the binomial, the exact form of the posterior is intractable, but the authors provide a normal approximation which facilitates inference. (p. 98)

Stiratelli, Laird and Ware (1984) introduced a different estimation procedure for a more general mixed model (similar to the Wong and Mason model) where they also assume that the logistic parameter for each group to be normally distributed in the population. Their estimation is based on the Maximum Likelihood estimation of fixed effects (the $\gamma$'s) and Maximum Likelihood using the EM algorithm for variance components and empirical Bayesian estimation of the random effects (the $U_{kj}$'s).

In fact, the approach of Stiratelli et al. is a generalization of Korn and Whittemore (1979) that assumes a logistic regression model with normally distributed random coefficients (i.e., random-effect model). Korn and Whittemore used a maximum likelihood estimation procedure that is based on a separate, logistic regression for each group. However, Korn and Whittemore's concern was to combine logistic regression coefficients across groups into a single logistic regression coefficient for each of the covariates (similar to Leonard, 1972b). Therefore, they did not include any group variables into their model.

Similarly, Clogg et al. (1990) introduced a simple Bayesian method in order to combine the logistic regression model across different regressions in a single equation. However, this estimation procedure is based on the maximum Posterior estimation that assumes Jeffrey's prior (i.e., noninformative prior, see Box and Tao, 1973 p.41; Rubin and Schenker, 1987) for the logistic regression model (i.e., B's).

Anderson and Aitkin (1985) derived a Maximum Likelihood estimation procedure in order to estimate the parameters in multilevel logistic and probit models. The logistic regression model was used, where the interviewee was considered as lower-level and the interviewer as upper-level of the hierarchy model. Their estimation is method based on the Bernoulli model for the binary response with the underlying assumption that the dependent variable is normally distributed. In addition, it was assumed that the random intercepts (random effect) had a normal distribution, and there was a fixed effects with its associated covariates.

Anderson and Aitkin concluded that the proportion of the variance of the dependent variable that is explained by variance component is nearly double the ANOVA estimate. They suggested that the use of ANOVA methods needs to be examined closely.

Longford (1988) used a quasi-likelihood estimation procedure based on the Nelder and Pregiborn (1987) estimation method. This is an extension of the quasi-likelihood estimation method "...to allow the comparison of variance function as well as those of linear predictors and link functions." (p. 221).

To obtain quasi-likelihood estimates, there is the need to define the quasi-likelihood function which is only to specify the relationship between the mean and the variance of the observation. But in order to define a likelihood function there is the need to specify the form of the distribution of the observation (Wedderburn, 1974). In fact, maximum quasi-likelihood estimates have many properties parallel to those of maximum likelihood estimates (Wedderburn, 1974; McCullagh, 1983).

Several assumption have been considered: (a) the usual assumptions of the normality of the random effects; (b) the non-normal error distribution; (c) the random effects, $U_{kj}$, $k=0,1,...,k$, are assumed to multivariate normal each with a mean of zero, and some variance, $Var(U_{kj}) = \tau_{kk}$. For any pair of random effects, $_k$ and $_{k'}$, $Cov(_k, _{k'}) = \tau_{kk'}$; (d) the assumption that the mean, $\theta_{ij}$, is related to linear predictors by a logit link function.

Thus, using "logit" as a link function will result in obtaining logistic regression coefficients having random slopes. Therefore, an estimate of within- and between-school parameters can be obtained (using the quasi-likelihood method) for a multilevel binary model (similar to the Wong and Mason model).

Goldstein (1989) proposed a multilevel nonlinear model when modelling discrete data. Here the within-group model for the binary outcome is specified with two dummy variables (k=2, $X_{1ij}$ and $X_{2ij}$). Thus the within-group logistic regression model is similar to Equation 2.6.

The between-group model assumes the $B_{kj}$ (k=2) to be random similar to Equation 2.5. Similar to other models. Goldstein assumed that; (a) the predictors are fixed, (b) the upper-level random terms $U_{oj}$ and $U_{1j}$ have a joint distribution with mean 0 and can be represented in a variance covariance matrix. As indicated for the within-group model, this model deals only with dummy variables for the within-group model by applying the iterative generalized least squares "IGLS" estimation method (Goldstein, 1986).

A real example of the above case that is provided in the ML3 manual (Prosser et al. 1991) is as follows: "...a sampled person working in factory j might be in one of eight job status categories, level 2 unit here are the factories, and level 1 units are the categories." (p. 22). Thus, the logit ($\theta_{ij}$) is considered as the dependent variable, where $\theta_{ij}$ is defined as the proportion of individuals in the job status category i in factory j that answered "Yes" to a "Yes" or "No" question.

Braun (1989) suggested a different estimation method for the hierarchical logistic regression model (i.e., the Wong and Mason model). Braun suggests first obtaining the ordinary logistic regression estimates (Equation 2.6), $\beta_i$ of $B_i$, along with the estimated $\sigma_i^2$ of these estimates ( $\beta_i \sim N(B_i,\sigma_i^2)$). Following this the empirical Bayesian estimates of $B_i$ can be derived from $\beta_i \sim N(B_i,\sigma_i^2)$ and the between-group model (Equation 2.5).

## Summary

Concerns regarding the distribution of normal errors in the case of data with binary outcomes have led to the development of the logistic regression model. However, there are some disadvantages of the logistic regression model when dealing with data sets involving two or more levels of hierarchy. For the past decade the concerns regarding the appropriate analysis of multilevel data structure have led to the studies of several methods of estimation for multilevel linear models with normally distributed outcomes.

However, several researchers have expressed concern regarding the use of multilevel linear model analysis when the normality assumption of the residuals is violated, specifically in the case of binary outcomes. This has led to development of several different approaches that take into account both the binary response and multilevel data structure (see Table 2-1). However, the popularity of these estimation methods and their applications in the field of educational research are limited.

# CHAPTER III

# METHOD

## Introduction

This chapter has been divided into three sections. The first section deals with the pilot study carried out on all four methods of estimation using the computer programs of ML3, VARCL, MULTILOGIT, and SPSS. These four computer programs will first be presented by describing the requirements for operating each program and indicating the initial advantages and disadvantages of each of them.

The second section will address the real data. A brief description of the real data will be presented. This will be followed by presenting the five multilevel logistic regression (MLR) models, and the two MLR models using the VARCL method of estimation.

The final section will address the simulated data. First, a description of the simulated model will be presented. Second, an account of the selected values for the conditions of interest that were used in the simulated model will be given. Third, the procedure used to generate simulated data will be described. Finally, the statistics used to evaluate the accuracy and properties of both VARCL and SPSS estimation methods will be presented.

## The Pilot Study

A pilot study was first conducted on the SPSS program which takes into account structure of the outcome for single-level data. Subsequently, a study was conducted on the other programs VARCL, ML3, MULTILOGIT which take into account the multilevel structure of the data.

A random sample of 20 schools was first drawn out of 411 schools from the real data (i.e., the Thailand data). A total of 406 students were found in the sample. Three dichotomous variables were also selected. These were (a) student repetition as dependent variable, (b) student sex as student-level covariate, and (c) school location (urban versus rural) as school-level (or group-level) covariate. Dichotomous variables were selected as covariates because the ML3 program requires that the two covariates be dichotomously coded.

The purpose of the pilot study was to run these four computer programs using the sample data in order to observe the advantages and disadvantages of these programs before conducting any further real or simulated data analysis. As such, the obtained estimates of these programs were not compared in this pilot study.

In the following account the researcher will introduce each of the computer programs, describe the requirements for operating each program, discuss the advantages and disadvantages for each program, and state the concerns of each program for further analysis in this study.

For simplicity, each estimation method will be identified by the name of the program. The programs will be identified as follows: the Maximum Likelihood estimation method as the SPSS program, the Quasi-Likelihood estimation method as the VARCL program, the Generalized Least Square estimation method as the ML3 program, and lastly, the Empirical Bayesian estimation method as the MULTILOGIT program.

## The VARCL Program

The VARCL program was first initiated by Aitkin and Longford (1986) and maintained by Longford. It is designed for the fitting of mixed linear models with nested random effects on data involving hierarchies of nesting.

The analysis using the VARCL program in this study was based on its microcomputer version. The researcher was also able to obtain the mainframe version of the program. It is useful to note that the interface of the VARCL program combines both an interactive and a batch feature of operation.

To run the VARCL program, the user needs to identify three input files, namely: (1) the basic information file, (2) the data file for student-level variables and the interaction term between student variables and school variables, and (3) the data file for school-level variables.

The following information should be furnished to the basic information file:

line 1: the research title,

line 2: the number of levels of nesting (two or three levels of nesting),

line 3: the number of units for both students and groups (schools),

line 4: the number of variables for both student- and school-levels,

line 5: the maximum number of iterations, frequency of report of convergence, and precision (a choice up to 4 decimal places, .0001),

line 6: the name of the unit-level (i.e., student), and name of the school-level (i.e., school),

line 7: the name of the file containing the student data,

line 8: the format of the student data,

line 9: the name of the file containing the school data,

line 10: the format of the school data.

The rest of the lines contain the name of the variables together with the number of its categories (this is equal to 1 for continuous variables), and finally, the number of subjects within each group. An example of the basic information file specified for this study is found in Appendix 3-1.

The specification of the model part and both the fixed and random effects of the model was done interactively.

The VARCL estimates converged to give the estimate of the parameters of the pilot data. The analysis of the results of the pilot revealed the following minor disadvantages of the VARCL program:

1) The model specification for the VARCL program is different from the MULTILOGIT program. The VARCL program macro (i.e., school variable) variable could not be specified as a predictor of the micro regression coefficient. However, the same MULTILOGIT and VARCL combined model can be obtained.
2) The independent covariates variables values had to be coded as "1"'s and "2"'s instead of "0"'s and "1"'s.
3) The user of the program has to specify the random effects twice to obtain the estimate of the random part of the model.

By running the VARCL program using the pilot data, it was found that the program uses standard logistic regression estimates (the same estimate up to four decimal places) as its initial estimate. However, the MULTILOGIT program requires that initial estimates be given for each specified model (within-school regression model) and for each school in the sample. It was also observed that the VARCL program converged to the estimates very rapidly. In addition, VARCL program was friendly and easy to use by combining both an interactive and a batch feature of operation. The manual for VARCL contained not only the information about the procedure to create an VARCL batch file and mixed model specification, but also provided many examples to assist the investigator.

## The ML3 Program

The ML3 software program is used for two and three-level multilevel data analysis by Rabash Prosser and Goldstein, based on Goldstein (1987). The researcher was able to obtain both the microcomputer and the mainframe version of the program. However, the analysis was based on the microcomputer version.

The ML3 program operates interactively. The user is required to identify a single data file that contains both the student and school variables, identifying each level by an identification code. The ML3 program is easy to use, and the furnish manual was sufficient, containing information about the estimation method, multilevel model specification, and procedure to operate the program.

Some major disadvantages of the ML3 program were revealed during the pilot analysis of the pilot data. ML3 requires that both the levels of the variables be dichotomous. In addition, ML3 requires the specification of the number of students ($n_{ij}$) in each sex by URB/RRL (urban or rural the school location variable) cell categories, and number of students from each of the $n_{ij}$ cells who repeated. These specific requirements made the running of the ML3 program very cumbersome.

In spite of detailed model specifications for the ML3 program design, the estimates of the ML3 programs in the pilot study did not converge. This may have been caused by having one urban school (with a total of 28 students) and 19 rural schools (with a total of 378 students) within the 20 schools randomly drawn.

Since the ML3 program requires dichotomous covariate variables at both levels and design specification, it was decided to drop out the ML3 estimation method from further data analysis. This is because both the design and covariates specifications would be different for real and simulated data. For example, the simulated independent variables for student and school level would be continuous variables, while the real data would contain both continuous and dichotomous variables. In addition, the ML3 program was comparatively much slower to run.

## The MULTILOGIT Program

The MULTILOGIT program was written by Albert F. Anderson, of the Population Studies Center at the University of Michigan, from instructions provided by George Y. Wong and William M. Mason. The program executes the multilevel logistic regression model that is proposed by Wong and Mason (1985).

The program is only available in the mainframe at the University of Michigan. To run the program, several manuals were required to explain how to operate the University of Michigan computer terminal system (MTS), and secondly how to use the MTS file editor. In addition, a PCTIE program needs to be purchased in order to allow the microcomputer to operate as a terminal to the University of Michigan network host. The PCTIE command also allows the transfer of files between the microcomputer and MTS.

In order to run the MULTILOGIT program, the user is required to specify four input files: (1) a micro (student) data input file, (2) a macro (school) data input file, (3) a coefficient input file, and (4) a command file.

These micro and macro data files contain the student- and school-level data sets. The coefficient input file contains the classical within-group (school) logistic regression coefficients for each school in the macro data. These coefficients will be used to generate starting values for the iterative algorithm. The command file performs the following functions: (1) It defines the multilevel logistic regression model, (2) provides terminating conditions for the algorithm, and (3) specifies input and output files (a copy of a command file is found in Appendix 3-2). The command file operates as a MULTILOGIT batch file.

Initially, there were some problems running the MULTILOGIT program, because the manual set-up specifications were a little different. When the program was finally run on the pilot data, estimates of the parameters were obtained.

Other than this initial problem, the program was easy to use, and the researcher had only to deal with the command file in order to change the multilevel logistic regression model. In addition, the program converged rapidly when used on the mainframe. The supplied manual contained sufficient information on how to write a command file, and run the program.

However, some limitations of the MULTILOGIT program were found and summarized as follows:

1. The micro (student) data file could only include 9 distinct micro variables (not counting the micro intercept).

2. The program could only read 5 macro variables (not counting the macro intercept).
3. The maximum number of schools (group) that could be used was 59.
4. The program required classical within-school logistic regression coefficients for each school (20 schools in the pilot study) in the analysis. These values had to be supplied by the researcher in order to generate the starting values for the iterative algorithm.
5. The MULTILOGIT program assumed that all micro regression (intercept and slopes) were random coefficients. In other words, the MULTILOGIT program did not accept the fixing of any within-school regression coefficient.
6. The model specification for the MULTILOGIT program was different from the VARCL program. However, the same combined model for the VARCL and the MULTILOGIT program was obtained.

The MULTILOGIT program specifies the school (or group) variables to be used as regressors in a between-school regression model in which the dependent is the slopes coefficient or intercepts. The difference between the model specifications of the VARCL and the MULTILOGIT program will be clarified later when the real data is analyzed using the different models for three estimation procedures, VARCL, MULTILOGIT and SPSS.

In addition to the above disadvantages, the cost of running the MULTILOGIT program on the mainframe computer of the University of Michigan was also a major financial concern.

### The SPSS Program

The SPSS program is a multi-purpose statistical package. It available on the mainframe at Michigan State University and also as a microcomputer version (both forms of SPSS were used in the analysis). The SPSS uses the single-level logistic regression model ignoring the hierarchical structure of the data. In other words, it assumes that the logistic regression parameter (slopes and intercepts) have fixed effects, ignoring the group (school) effect on the variability between slopes and intercepts. This model specification is considered as the major disadvantage of this program.

In the SPSS program the maximum-likelihood method of estimation is used to obtain the estimates of the logistic regression model parameters. In addition, since the model is nonlinear, an iterative algorithm is used for parameter estimation.

## The Finding of the Pilot Study

The pilot study revealed several limitations, advantages and disadvantages of the four computer programs.

The ML3 program requires dichotomous covariate variables at both levels and has an inconvenient design specification. Because of this, a decision was made to exclude the ML3 estimation method from further data analysis. The maximum number of schools (group) that can be used (59) was a serious limitation of the MULTILOGIT program.

In addition, the MULTILOGIT program requires the specification of the classical within-school logistic regression coefficients for each of the schools in the analysis. This proved to be too cumbersome. Furthermore, the MULTILOGIT program always assumes that the micro regression coefficients are random. Finally, the cost of running the MULTILOGIT program on the mainframe computer at the University of Michigan proved to be a major financial concern especially when considered for use with simulated data.

In fact, the cost of running the MULTILOGIT program at the University of Michigan mainframe computer center lead the researcher to run VARCL, SPSS, and MULTILOGIT on the real data first rather than the simulated data in order to determine the real cost. This allowed the researcher to predict the extremely high financial cost of running the MULTILOGIT program on the simulated data.

## Characteristics of the Real Data

The real data analyses were based on data from Thailand collected in 1988 under the sponsorship of the BRIDGES (Basic Research in Developing Educational Systems) project. A random sample of 59 schools (due to the limitation of the MULTILOGIT program on the maximum number of groups) consisting of 1244 sixth-grade students was utilized. The analysis was based on several models, from simple to more complex, using two student variables: (1) the student repetition where "1" indicates "ever" and "0" indicates "never", and (2) student socioeconomic status (i.e., SES).

In addition to this, five school-level variables were also included: (1) the school location (urban vs. rural), (2) school SES (i.e., MEAN SES, the student SES was aggregated at the school level to measure the school SES), and (3) three geographic variables. These were allocated in terms of location of the school in the central, north or south regions of Thailand.

These variables were chosen based on previous work which indicated that they were related to student repetition. The descriptive statistics for the real data at both the student- and the school-level are presented in Appendix 3-3.

All the variables at both levels had to be centered in order to be able to compare the regression coefficients across the different approaches. This was because the VARCL program centered all the variables.

Each of the three programs were run several times in order to ensure that the data loaded on to the program had been read accurately.

## Multilevel Logistic Regression Models

Five different Multilevel Logistic Regression (MLR) Models, from simple to complex were analyzed. This was done in order to compare the following estimated statistics in the three estimation procedures: (a) the estimates of the regression coefficients and their standard errors, and (b) the variance and covariance of the random effects and their standard errors.

An additional analysis was also performed using the VARCL program comparing the random intercept and fixed regression coefficient model with the combined random intercept and random regression coefficient model. This was done in order to show advantages of using one model over the other, and the ability of the VARCL program to test the variance and covariance of the random effects of the model (i.e., $H_o: \tau_{oo}=0$, $H_o: \tau_{11}=0$, and $H_o: \tau_{o1}=0$).

Each of these five multilevel logistic regression models will be presented in this chapter. The results of the real data analysis and the comparisons between the estimated statistics, using three methods of estimation, will be subsequently presented in chapter four.

## MLR Model 1

The simplest MLR model considered in this study included no student-level and school-level independent variable as covariates.

The within-school equation (or group) for MULTILOGIT and VARCL is represented as

$$\text{Logit (repetition)}_{ij} = B_{oj} \quad (3.1)$$

The between-school equation (or group) for MULTILOGIT and VARCL is represented as

$$B_{oj} = \gamma_{oo} + U_{oj} \quad (3.2)$$

where

$B_{oj}$ is the average of the Logit (repetition)$_{ij}$ in school j (Equation 3.2 shows that $B_{oj}$ varies around the grand mean $\gamma_{oo}$ with variance $\sigma^2(U_{oj})=\tau_{oo}$).

$U_{oj}$ is the random effect associated with school j.

Thus the combined equation of MLR model 1 for MULTILOGIT and VARCL was obtained by substituting Equation 3.2 into 3.1,

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + U_{oj} \quad (3.3)$$

The MLR model 1 for SPSS is simply represented as,

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} \quad (3.4)$$

MLR model 1 for MULTILOGIT and VARCL is a useful way to estimate much of the variation that exists in the dependent variable between schools. It is clear that the MLR model for SPSS does not account for the between-school (or group) variation (compare Equation 3.3 with 3.4).

<u>MLR Model 2</u>

The second MLR model considered in this study included the student-level variable of student socioeconomic status (i.e., SES) in the within-school equation as a covariate. The school-level variable was also excluded.

The within-school equation (or group) for MULTILOGIT and VARCL is represented as

$$\text{Logit (repetition)}_{ij} = B_{oj} + B_{1j}\,(\text{SES})_{ij} \quad (3.5)$$

The between-school equations (or group) for MULTILOGIT and VARCL is represented as

$$B_{oj} = \gamma_{oo} + U_{oj} \tag{3.6}$$

$$B_{1j} = \gamma_{1o} + U_{1j} \tag{3.7}$$

where

$B_{oj}$ is the adjusted school mean (i.e., the raw school mean minus an adjustment for its SES mean); and

$B_{1j}$ is the effect of the student SES on the outcome within school j.

In the case above, both the adjusted school mean, $B_{oj}$, and the school regression coefficients, $B_{1j}$, vary across schools around their grand mean.

Thus the combined equation of MLR model 2 for MULTILOGIT and VARCL was obtained by substituting Equation 3.6 and 3.7 into 3.5,

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + \gamma_{1o}\ (SES)_{ij} + U_{1j}\ (SES)_{ij} + U_{oj} \tag{3.8}$$

The error term in Equation 3.8 is presented as $(U_{1j}\ (SES)_{ij} + U_{oj})$.

The MLR model 2 for SPSS is simply represented as

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + \gamma_{1o}\ (SES)_{ij} \tag{3.9}$$

<u>MLR Model 3</u>

The third MLR model considered in this study was similar to the second model. The only difference was that the school-level variable (i.e., school SES, MSES) was included as a covariate.

This model was specified differently for the MULTILOGIT and the VARCL programs. However, the combined MLR model 3 for both programs was identical.

The within-school equation for MULTILOGIT is similar to Equation 3.5.

The between-school equations for MULTILOGIT is represented as

$$B_{oj} = \gamma_{oo} + \gamma_{o1} \text{ (MSES)}_j + U_{oj} \qquad (3.10)$$

$$B_{1j} = \gamma_{1o} + U_{1j} \qquad (3.11)$$

The within-school equation for VARCL is represented as

$$\text{Logit (repetition)}_{ij} = B_{oj} + B_{1j} \text{ (SES)}_{ij} + \gamma_{o1} \text{ (MSES)}_j \qquad (3.12)$$

The between-school equations for VARCL model 3 is similar to Equation 3.6 and 3.7.

Thus, the combined MLR model 3 for MULTILOGIT (substituting Equation 3.10 and 3.11 into 3.5), and VARCL (substituting Equation 3.6 and 3.7 into 3.12) is derived as

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + \gamma_{1o} \text{ (SES)}_{ij} + \gamma_{o1} \text{ (MSES)}_j + U_{1j} \text{ (SES)}_{ij} + U_{oj} \qquad (3.13)$$

The MLR model 3 for SPSS is represented as

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + \gamma_{1o} \text{ (SES)}_{ij} + \gamma_{o1} \text{ (MSES)}_j \qquad (3.14)$$

MLR Model 4

In the fourth MLR model, another school-level variable (i.e., school location, urban versus rural, URB/RRL) was included as a covariate regressed on the regression slopes (i.e., $B_{1j}$) only for the within-school model. This variable was added in order to compare the interaction coefficient that was associated with it (i.e., $(\text{URB/RRL})_j * (\text{SES})_{ij}$). Thus, only the result associated with the URB/RRL variable will be discussed.

The within-school equation for MULTILOGIT is similar to Equation 3.5.

The first between-school equation for MULTILOGIT, (associated with the random intercept) is similar to Equation 3.10, while the second between-school equation (associated with SES regression slope) is represented as

$$B_{1j} = \gamma_{1o} + \gamma_{11} \ (URB/RRL)_j + U_{1j} \qquad (3.15)$$

The within-school equation for VARCL is represented as

$$\text{Logit (repetition)}_{ij} = B_{oj} + B_{1j} \ (SES)_{ij} + \gamma_{o1} \ (MSES)_j + \gamma_{11} \ ((URB/RRL)_j * (SES)_{ij}) \qquad (3.16)$$

The between-school equation for VARCL, is similar to Equation 3.6 and 3.7.

Thus, the combined MLR model 4 for MULTILOGIT (substituting Equation 3.15 and 3.10 into 3.5), and VARCL (substituting Equation 3.6 and 3.7 into 3.16) is derived as

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + \gamma_{1o} \ (SES)_{ij} + \gamma_{o1} \ (MSES)_j + \gamma_{11} \ ((URB/RRL)_j * (SES)_{ij}) + U_{1j} \ (SES)_{ij} + U_{oj} \qquad (3.17)$$

The MLR model 4 for SPSS is represented as

$$\text{Logit (repetition)}_{ij} = \gamma_{oo} + \gamma_{1o} \ (SES)_{ij} + \gamma_{o1} \ (MSES)_j + \gamma_{11} \ ((URB/RRL)_j * (SES)_{ij}) \qquad (3.18)$$

MLR Model 5

In this model the school-level variables for geographical region were included.

The within-school equation for MULTILOGIT, is similar to Equation 3.5.

The first between-school equation for MULTILOGIT is represented as

$$B_{oj} = \gamma_{oo} + \gamma_{o1}\ (URB/RRL)_j + \gamma_{o2}\ (CENTRAL)_j + \gamma_{o3}\ (NORTH)_j + \gamma_{o4}\ (SOUTH)_j + \gamma_{o5}\ (MSES)_j + U_{oj} \quad (3.19)$$

While, the second between-school equation (associated with SES regression slope) is similar to Equation 3.15. The within-school equation for VARCL is represented as

$$Logit\ (repetition)_{ij} = B_{oj} + B_{1j}\ (SES)_{ij} + \gamma_{o1}\ (URB/RRL)_j + \gamma_{o2}\ (CENTRAL)_j + \gamma_{o3}\ (NORTH)_j + \gamma_{o4}\ (SOUTH)_j + \gamma_{o5}\ (MSES)_j + \gamma_{11}\ ((URB/RRL)_j * (SES)_{ij}) \quad (3.20)$$

The between-school equation for VARCL, is similar to Equation 3.6 and 3.7.

Thus, the combined MLR model 5 for MULTILOGIT (substituting Equation 3.15 and 3.19 into 3.5) and VARCL (substituting Equation 3.6 and 3.7 into 3.20) is derived as

$$Logit\ (repetition)_{ij} = \gamma_{oo} + \gamma_{1o}\ (SES)_{ij} + \gamma_{o1}\ (URB/RRL)_j + \gamma_{o2}\ (CENTRAL)_j + \gamma_{o3}\ (NORTH)_j + \gamma_{o4}\ (SOUTH)_j + \gamma_{o5}\ (MSES)_j + \gamma_{11}\ ((URB/RRL)_j * (SES)_{ij}) + U_{1j}\ (SES)_{ij} + U_{oj} \quad (3.21)$$

The MLR model 5 for SPSS is represented as

$$Logit\ (repetition)_{ij} = \gamma_{oo} + \gamma_{1o}\ (SES)_{ij} + \gamma_{o1}\ (URB/RRL)_j + \gamma_{o2}\ (CENTRAL)_j + \gamma_{o3}\ (NORTH)_j + \gamma_{o4}\ (SOUTH)_j + \gamma_{o5}\ (MSES)_j + \gamma_{11}\ ((URB/RRL)_j * (SES)_{ij}) \quad (3.22)$$

## Comparing Two MLR Models Using the VARCL Method of Estimation

This analysis will show the advantages of using one model over the other. This comparison is made possible because of the VARCL program's ability to test the variance and covariance of the random effects. Two models, A and B, will be specified in this study. Model A having a random intercept and random regression slope, and model B having a random intercept model and fixed regression slope.

For model A the within-school, between-school, and the combined VARCL equation are the same as the VARCL MLR model 5 (refer to equations 3.20, 3.6, 3.7, and 3.21).

For model B the within-school equation is similar to Equation 3.20. The first between-school equation (associated with the random intercept) is similar to Equation 3.6, while the second between-school equation (associated with the SES regression slope) is represented as

$$B_{1j} = \gamma_{1o} \tag{3.23}$$

Thus, the combined model B for VARCL (substituting Equation 3.23 and 3.6 into 3.20) is derived as

$$\begin{aligned} \text{Logit (repetition)}_{ij} &= \gamma_{oo} + \gamma_{1o}\ (\text{SES})_{ij} + \gamma_{o1}\ (\text{URB/RRL})_j \\ &+ \gamma_{o2}\ (\text{CENTRAL})_j + \gamma_{o3}\ (\text{NORTH})_j + \gamma_{o4}\ (\text{SOUTH})_j + \gamma_{o5}\ (\text{MSES})_j \\ &+ \gamma_{11}\ ((\text{URB/RRL})_j * (\text{SES})_{ij}) + U_{oj} \qquad (3.24) \end{aligned}$$

The only difference between the combined model A (i.e., Equation 3.21), and combined model B (i.e., Equation 3.24) is that model B suppresses the error term associated with SES, $U_{1j}$ $(\text{SES})_{ij}$.

The results of the analysis running the SPSS, VARCL and MULTILOGIT estimation methods on the five proposed multilevel logistic regression models and two proposed MLR models (A and B) using the VARCL method of estimation using real data (Thailand data) will be presented in chapter four.

Rationale for Excluding the MULTILOGIT Program

A more complicated multilevel logistic regression model was attempted by including more covariates in the within-school model. The results of running this new model were obtained for both the SPSS and VARCL methods of estimation.

However, the MULTILOGIT program did not run with this new model. It registered "bomb out" indicating an error message. The investigation into why this occurred revealed yet another disadvantage of the MULTILOGIT program. In order for the MULTILOGIT program to run, a specification of the coefficient input file is required. This file contains the classical within-group (school) logistic regression coefficients that will be used to generate the starting values for the iterative algorithm. These regression coefficients are obtained by estimating the classical logistic regression coefficients separately for each group (school) using the SPSS program. Thus, for each of the five specified models in this study, the logistic regression coefficient for each of the 59 schools was obtained.

This meant that each data line in the coefficient input file of MULTILOGIT was associated with a single school (59 different data lines for the 59 schools in each model) containing the intercept and the regression slope of the within-group logistic regression model.

However, since the number of students in each of the 59 schools range from a minimum of 8 to a maximum of 37 students, the within-school logistic regression coefficient (intercepts and slopes) estimates for the MULTILOGIT coefficient input file were estimated as zero. This caused the MULTILOGIT program to "bomb out". In fact, the concern regarding the number of subjects within each group was also mentioned in Wong and Mason (1985).

Based on the results of the pilot study and real data analysis, it was decided to exclude the MULTILOGIT program method of estimation from the simulated data analysis. This decision was based on the following reasons:

1. The high financial cost of running the MULTILOGIT program on the University of Michigan Mainframe Computer Center. Despite running the program in the minimum charge time, which was generally between 2:00 a.m. to 7:00 a.m., the estimated cost of running the MULTILOGIT program on 1200 replications simulated data would be at least US$15,000.00. This figure was based on the cost of running the MULTILOGIT program on the real data and pilot data of this study.
2. The MULTILOGIT program will not run in 600 out of 1200 replications of the simulated data. This is because the simulated condition for the number of subjects (students) within each group is considered as 10 (n=10). This will cause the within-group (school) logistic regression coefficients (intercepts and slopes) data for the MULTILOGIT coefficient input file to be estimated as zero.

3. For each of 1200 replications of the simulated data, 60 classical (i.e., SPSS estimates) within-group logistic regression coefficients need to be specified. These sixty data line estimates of the logistic regression coefficients for each replication point in the MULTILOGIT coefficient input file are due to the number of groups within each simulated replication. Obtaining all the coefficients would entail an enormous task.

As a result of this, the analysis of the simulated data was conducted using only the two methods of estimation: (a) VARCL, designed for data involving hierarchies of nesting having a binary outcomes, and (b) SPSS, designed for single-level model having binary outcomes.

## Characteristics of the Simulated Model

The simulated model was a two-stage multilevel logistic regression model having random intercept and a random regression coefficient.

The with-group model is represented as

$$\alpha_{ij} = B_{oj} + B_{1j} X_{ij} \qquad (3.25)$$

The between-group model is represented as

$$B_{oj} = \gamma_{oo} + \gamma_{o1} Z_j + U_{oj} \qquad (3.26)$$

$$B_{1j} = \gamma_{1o} + \gamma_{11} Z_j + U_{1j} \qquad (3.27)$$

The generated data has, within each group (or school), the micro predictor, $X_{ij}$, normally distributed with mean of zero and a variance of one (i.e., $X_{ij}$~N(0,1)). Similarly, the macro predictor, $Z_j$, is normally distributed with mean of zero and a variance of one (i.e. $Z_j \sim N(0,1)$). In addition, $U_{oj}$ and $U_{1j}$ are mutually independent, as they are generated separately (i.e., $\tau_{o1} = 0$).

The random effects (i.e., $U_{oj}$ and $U_{1j}$) were generated having both a normal distribution (ND) and t-distribution (TD). The normal distribution of the random effects were investigated under both a large magnitude of the random regression slope variance, 17.6% (denoted by RRSL), and a small magnitude of the random regression slope variance .005% (denoted by RRSS) of the intercept variance. Therefore, the random effects with RRSL were generated having a normal distribution with mean of zeros and variance-covariance components as

$$\text{Var}\begin{bmatrix} U0j \\ U1j \end{bmatrix} = \begin{bmatrix} \tau oo & \tau o1 \\ \tau 1o & \tau 11 \end{bmatrix} = \begin{bmatrix} .85 & .00 \\ .00 & .15 \end{bmatrix}$$

The RRSS the random effects were also generated to have a normal distribution with mean of zeros and variance-covariance components as

$$\text{Var}\begin{bmatrix} U0j \\ U1j \end{bmatrix} = \begin{bmatrix} \tau oo & \tau o1 \\ \tau 1o & \tau 11 \end{bmatrix} = \begin{bmatrix} .995 & .00 \\ .00 & .005 \end{bmatrix}$$

While, the t-distribution of the random effects were investigated only under a large magnitude of the random regression slope variance of the intercept variance. Thus the random effects with RRSL were generated having a t-distribution with four degrees of freedom with mean of zeros and variance-covariance components as

$$\mathrm{Var}\begin{bmatrix} U0j \\ U1j \end{bmatrix} = \begin{bmatrix} \tau oo & \tau o1 \\ \tau 1o & \tau 11 \end{bmatrix} = \begin{bmatrix} .85 & .00 \\ .00 & .15 \end{bmatrix}$$

In addition, the following values were chosen for $\gamma$'s: $\gamma_{oo}$ = -1.80, $\gamma_{o1}$ = -1.20, $\gamma_{1o}$ = -.50, and $\gamma_{11}$ = .75 . These values were chosen in order to represent realistic values of the situation. In fact, these values for $\gamma$'s were obtained from the previous analysis of real data (i.e., the results of the fourth model of the real data analysis using three main estimation methods). Similarly, the simulated true values for the variance-covariance components of random effects (i.e., $\tau_{oo}$ and $\tau_{11}$). Thus by substituting the $\gamma$'s values into Equation 3.26 and 3.27, and further substituting the two equations (i.e., 3.26, and 3.27) into equation 3.25 the combined model is derived as

$$\alpha_{ij} = -1.8 - 1.20\ Z_j - .50\ X_{ij} + .75\ (X_{ij} * Z_j) + (U_{1j}\ X_{ij} + U_{oj}) \quad (3.28)$$

The accuracy and properties of the parameter estimation for both VARCL and SPSS programs of the simulated data were evaluated under the moderate number of sixty schools, j=1,2,...60.

In addition, the statistical properties of these estimation procedures were investigated under realistic values for the three following conditions:

1. Number of subjects within each group (n).

   Simulating data with n = 10 and 60 units (subjects) within each group. These small and large values of the number of subjects within group was based on several studies (Bock, 1983; Aitken and Longford, 1986; Wong and Mason, 1985).

2. Magnitude of the random regression slope (RRS) variance.

   Specifying the RRS variance of .005% (small variance denoted by RRSS) and 17.6% (Large variance denoted by RRSL) of the intercept variance, the following values were chosen, $\tau_{00} = .995$ and $\tau_{11} = .005$ to obtain RRSS and $\tau_{00} = .85$ and $\tau_{11} = .15$ to obtain RRSL. These values were selected based on (Wong and Mason, 1985) and the results of the previous analysis of the real data.

3. Normal distribution (ND) and t-distribution (TD) of the random effects, $U_{0j}$ and $U_{1j}$.

## The Design of the Simulated Study

In order to establish the design of the simulated study, three conditions of interest (with two levels within each condition) had to be considered. This resulted in a design that consisted of a total of six cells see Figure 3-1.

For simplicity, each cell was identified by the following notations:

**(ND,n10,RRSS)** defining a normal distribution (ND) of the random effects, with 10 (signifying a small number) subjects within each cell (n10), and small random regression slope (RRSS) of the intercept variance.

**(ND,n10,RRSL)** defining a normal distribution (ND) of the random effects, with 10 subjects within each cell (n10), and large random regression slope (RRSL) of the intercept variance.

**(ND,n60,RRSS)** defining a normal distribution (ND) of the random effects, with 60 (signifying a large number) subjects within each cell (n60), and small random regression slope (RRSS) of the intercept variance.

**(ND,n60,RRSL)** defining a normal distribution (ND) of the random effects, with 60 subjects within each cell (n60), and large random regression slope (RRSL) of the intercept variance.

**(TD,n10,RRSL)** defining a t-distribution (TD) of the random effects, with 10 subjects within each cell (n10), and large random regression slope (RRSL) of the intercept variance.

**(TD,n60,RRSL)** defining a t-distribution (TD) of the random effects, with 60 subjects within each cell (n60), and large random regression slope (RRSL) of the intercept variance.

<u>Procedure Used to Generate the Simulated Data</u>

A Gauss computer program was used with an IBM compatible 386/Mhz microcomputer to generate the data for each of the six cells (combing the three conditions of interest and equations satisfying 3.25, 3.26 and 3.27). A math coprocessor was installed in the microcomputer to speed up the process.

Since the analysis is based on a moderate number of 60 groups, a vector of 60 by 1 was first generated for the group predictor, $Z_j$, having a normal distribution with a mean of zero and a variance of one. A copy of the program is shown in Appendix 3-4.

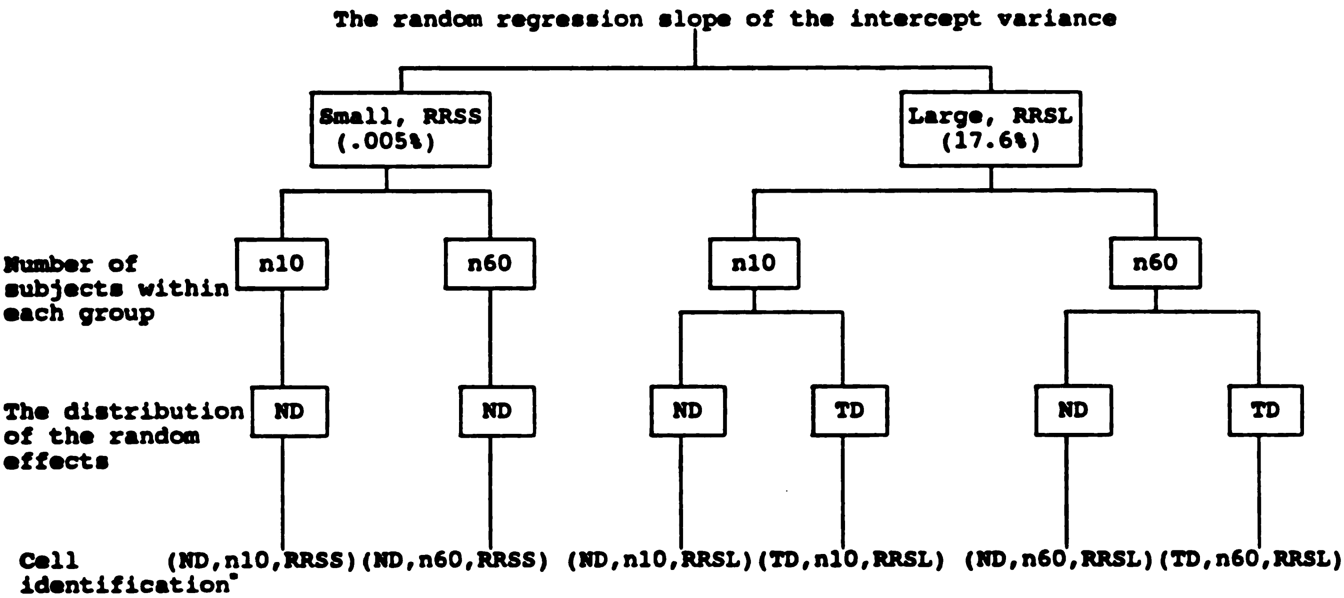
In addition, a 600 by 1 vector was generated for the within-group predictor, $X_{ij}$. This was because there were 10 subjects within each of the 60 groups resulting in 10 by 60 (i.e., 600) subjects of $X_{ij}$ covarite being generated, having a normal distribution within each group (or school) with a mean of zero and a variance of one. A copy of the program is found in Appendix 3-5.

Similarly, a 3600 by 1 vector was also generated for $X_{ij}$ and used where the number of subjects within each group was taken to be sixty, n=60. Note that both $Z_j$ and $X_{ij}$ are considered as fixed variables.

The random effects, $U_{oj}$ and $U_{1j}$, of the equations 3.26 and 3.27 were also generated satisfying the conditions of interest (i.e., RRSS or RRSL, and ND or TD).

The random effects, $U_{oj}$ and $U_{1j}$, were generated having a t-distribution with four degrees of freedom. Selecting four

Figure 3-1.--The design of the six cells of the simulated study




* 200 replication within each cell.

degrees of freedom would make the t-distribution deviate from the normal distribution. The distribution of the random effects was checked in order to ensure that the program was working properly.

Thus, using the generated $U_{oj}$, $U_{1j}$, $Z_j$ and the assigned values of $\gamma$'s ($\gamma_{oo} = -1.80$, $\gamma_{o1} = -1.20$, $\gamma_{1o} = -.50$, and $\gamma_{11} = .75$), together with equations 3.26 and 3.27, the $B_{oj}$'s and $B_{1j}$'s were computed for each group.

Using the computed $B_{oj}$'s, $B_{1j}$'s for each group and the generated $X_{ij}$ values for each subject within the same group derived from Equation 3.25, $\alpha_{ij}$ was computed for each unit (i.e., the individual subject) within each group (school).

Since the objective of the study was to obtain binary (0's and 1's) outcomes for individual units, $\theta_{ij}$ was first obtained from $\alpha_{ij}$ by using the procedure $\theta_{ij} = e^{\alpha ij} / (1 + e^{\alpha ij})$, ($\theta_{ij} = P(Y_{ij}=1)$ is the probability that the i-th micro observation will select the first category (i.e., $Y_{ij}=1$) of the response variable. This was obtained by solving the equation, below.

Since $$\mathrm{Ln}\,[\theta_{ij} / (1- \theta_{ij})] = \alpha_{ij}$$

$$\theta_{ij} / (1- \theta_{ij}) = e^{\alpha ij}$$

$$\theta_{ij} = (1- \theta_{ij})\, e^{\alpha ij}$$

$$\theta_{ij} = e^{\alpha ij} - \theta_{ij}\, e^{\alpha ij}$$

$$\theta_{ij} + \theta_{ij}\, e^{\alpha ij} = e^{\alpha ij}$$

$$\theta_{ij}\, (1 + e^{\alpha ij}) = e^{\alpha ij}$$

Thus $$\theta_{ij} = e^{\alpha ij} / (1 + e^{\alpha ij}) = 1 / (1 + e^{-\alpha ij})$$

Finally, assuming that $Y_{ij}|\theta_{ij} \sim \mathrm{Bernoulli}\ (\theta_{ij})$, the binary scores for each unit (individual student outcome), $Y_{ij}$, was obtained.

By drawing a number at random from a uniform distribution within the range of zero to one, $Y_{ij}$ is assigned to value one, if the value of the random number is less than or equal to the $\theta_{ij}$. If the random number drawn exceeds $\theta_{ij}$, $Y_{ij}$ is recorded as zero.

Each of the above steps in this simulated study was checked thoroughly in order to be confident that the simulation program was doing what it was expected to do.

For each of the six cells, 200 replications were performed. However, due the space limitation in the hard disk of the researcher's microcomputer, the program was run twice with 100 replications each time in order to obtain the 200 replications for each cell.

A total of 1200 replications were performed (200 replications for each of the six cells), generating two sets of data that could be used on the SPSS and VARCL programs.

Fortunately, it was discovered from the real data and the simulated data analysis that the VARCL program used the SPSS estimates (the same values, up to four decimal places) as its initial values in order to obtain VARCL estimates. Thus, only one set of generated data was applied to VARCL program as the SPSS estimates could be obtained from the VARCL printout.

The data obtained from each replication run were,

$Y_{ij}$ unit outcomes (the dependent variable for each unit or student within a group),

$X_{ij}$ unit within group predictor,

$Z_j$ group predictor, and

$X_{ij}*Z_j$ the interaction term between the within-group predictor and the group-level predictor.

A copy of the GAUSS program that was used to generate a data set for the cell (ND,n10,RRSS) is found in Appendix 3-6.

In addition, the statistics, $\tau_{00}$, $\tau_{11}$ and $\tau_{01}$, for each replication were saved on separate files. This was done in order to compare their standard errors with the estimated standard error of variance-covariance of the random effects from the VARCL estimation procedure later.

The VARCL program was run on each set of the 1200 replications (data set) resulting in a printout of 1200 values consisting of the SPSS and VARCL estimates. These parameter estimates of the simulated model were used to evaluate both the estimation methods.

The estimated parameters that were obtained for both the SPSS and VARCL procedures were saved on a single file. These statistics were later used to evaluate the accuracy and properties of both the VARCL and SPSS estimation procedures.

The average time to obtain 100 replications (100 sets of data for the VARCL program) for n=10 was approximately 45 minutes and approximately 2 hours for n=60. The average time of running the VARCL program for one data set, where n=10, was approximately 1.4 minutes and approximately 4 minutes for n=60.

## Statistical Comparison of the Estimation Methods

To compare the two estimation methods the following statistics were computed :

1. The mean (average) of estimates, $E(\hat{\gamma}_i) = \bar{\hat{\gamma}}$ where $\hat{\gamma}_i$ is the VARCL and SPSS estimate for each of the parameters of interest.
2. The bias, $E(\hat{\gamma}_i - \gamma_t)$ where $\gamma_t$ is the true parameter value. A 95% confidence interval of the bias (bias ± 1.96 x S.E.(bias)) was also constructed. In addition, the magnitude of the bias was compared to the true value. The

percent of this bias is obtained by dividing the absolute value of the bias by the absolute true value and multiplying the value obtained by 100.

3. The mean square error (MSE) of estimates. This statistic combined the bias and the dispersion of an estimator into a single quantity:

   $$MSE = E(\hat{\gamma}_i - \gamma_t)^2 = VAR(\hat{\gamma}) + E(\hat{\gamma}_i - \gamma_t)^2 = VAR(\hat{\gamma}) + Bias^2$$

4. The probability of type I error rate under a true null hypothesis ($H_o$: $\gamma = \gamma_t$ ). This is determined by counting the frequency with which the test statistic, $Z = (\hat{\gamma} - \gamma_t)/S.E.(\hat{\gamma})$ in each replication, exceeds a specified critical value (at .05 significance level), and dividing by the total number of replications.

The results of the analysis comparing the VARCL and SPSS estimation methods using simulated data will be presented in Chapter Four.

## Summary

The research study began with a pilot study in order to evaluate the multilevel and single-level logistic regression models as analyzed by four computer programs: ML3, VARCL, MULTILOGIT, and SPSS. As a result of this, the ML3 program was excluded from the analysis, and some concerns arose regarding the MULTILOGIT program. The computer programs were then run using the real data and a simulation exercise was also executed. Five multilevel logistic regression models and two MLR models using the VARCL method of estimation were demonstrated for the real data analysis. As a result of the real data analysis, the MULTILOGIT program was excluded from the study. The procedure for running the simulation study using the GAUSS program was also explained.

# CHAPTER IV

# RESULTS

## Introduction

The results of this study are presented in three sections. The chapter begins by presenting the results of the real data analysis. The second section will address the results of the simulation data analysis, comparison of the SPSS and VARCL estimation methods, and evaluation of the effect of the three conditions on both estimation procedures. The third section will deal with the accuracy of the VARCL estimation method and the properties of the estimates of the variance-covariance components of random effects and its standard errors with the true values.

## Results of the Real Data Analysis

The results of running the SPSS, VARCL, and MULTILOGIT programs of estimation methods of the real data will be presented for each of the five specified multilevel logistic regression (MLR) models and the two MLR models using the VARCL estimation procedure (i.e., model A versus model B).

### The Results of MLR Model 1

The results of the analysis indicate that the absolute value of intercept coefficient, $\gamma_{00}$, for the MULTILOGIT approach had the largest value (-1.92074) followed by the SPSS and VARCL approach with values of -1.7010 and -1.632833, respectively (See Table 4-1a). The difference of the $\gamma_{00}$ estimate between the SPSS and VARCL estimation methods was very small.

Table 4-1a.--Estimated regression coefficient and standard error (given in parentheses) for MLR model 1 using different estimation methods.

| | Estimation Method | | |
|---|---|---|---|
| | M.L. (SPSS) | Quasi-likelihood (VARCL) | Empirical Bayesian (MULTILOGIT) |
| Intercept, $\gamma_{00}$ | -1.7010 (.0785) | -1.632833 (NA) | -1.92074 (.1770906) |

Note: NA-not given by the program.

In addition, the results indicate that the MULTILOGIT estimate of the standard error for $\gamma_{00}$ (intercept coefficient) is larger (.17709) than the SPSS estimate (.0785), while the VARCL program did not report the standard error for $\gamma_{00}$. This is one disadvantage of using the VARCL program.

The variance of the intercept of the random effects, $\tau_{00}$, was also compared using the MULTILOGIT and VARCL methods (the SPSS method did not report $\tau_{00}$ since the program does not account for between-school variation). The results shown in Table 4-1b indicate a slightly larger variance of the intercept of the random effects for MULTILOGIT (1.29436) than for VARCL (1.084792). This is because VARCL program uses approximation of the maximum likelihood estimate.

Table 4-1b.--Estimated variance of the intercept random effects and S.E. for $\sqrt{\tau_{00}}$ (given in parentheses) for MLR model 1 using the VARCL and MULTILOGIT estimation methods.

| | Estimation Method | |
|---|---|---|
| | Quasi-likelihood (VARCL) | Empirical Bayesian (MULTILOGIT) |
| Intercept, $\tau_{00}$ | 1.084792* (.128574) | 1.29436 (NA) |

Note
(i) NA-not given by the program.
(ii) *-significant at .05 ($H_o$: $\tau_{00}$=0, t-statistic =1.0415/.1286=8.1007).

The VARCL program also reported a standard error for the intercept standard deviation of the random effects, .128574 (standard error of $\sqrt{\tau_{00}}$). This provided a significance test of the between-school variation. The null hypothesis here was: $H_o$: $\tau_{00}=0$. The t-statistic test of significance at p = .05 indicated that the null hypothesis should be rejected and that there were significant differences among schools with respect to their mean outcomes (i.e. Logit $(\text{repetition})_{ij}$), t=1.04153/.128574 = 8.1007 ($\sqrt{\tau_{00}}$).

The hypotheses testing of variance-covariance components (i.e., $H_o:\tau_{00}=0$, $H_o:\tau_{11}=0$ and $H_o:\tau_{01}=0$) of the VARCL results help investigators decide whether the regression intercepts and slopes for the within-group model should be specified as fixed or random, only when $\tau$ is significant. Note that the SPSS estimation method assumed that the regression intercepts and slopes were fixed, thus ignoring the variation of these regression coefficients among the groups.

By testing the above hypotheses the researcher is able to make a proper decision as to whether the regression coefficient among schools is fixed or random. This testing feature is found only in the VARCL program.

## The Results of MLR Model 2

The results (Table 4-2a) indicate that both the intercept coefficient, $\gamma_{oo}$, and its standard error (S.E.($\gamma_{oo}$)) have a pattern similar to the MLR model 1.

Table 4-2a.--Estimated regression coefficient and standard error (given in parentheses) for MLR model 2 using different estimation methods.

| | Estimation Method | | |
|---|---|---|---|
| | M.L. (SPSS) | Quasi-likelihood (VARCL) | Empirical Bayesian (MULTILOGIT) |
| Intercept, $\gamma_{oo}$ | -1.8063<br>(.0873) | -1.750708<br>(NA) | -1.98894<br>(.1687909) |
| SES, $\gamma_{1o}$ | -.8841<br>(.1726) | -0.545613<br>(0.197848) | -0.645738<br>(.2290343) |

Note: NA-not reported by the program.

Comparing the slope regression coefficients associated with SES, $\gamma_{1o}$, the MULTILOGIT and VARCL estimates of $\gamma_{1o}$ were found to be (-0.6457) and (-0.5456), respectively. This indicates that they were quite close in value to each other, and the absolute value of MULTILOGIT estimate is slightly larger than VARCL estimate. While, the SPSS estimate was somewhat larger (-.8841).

The results also indicate that the MULTILOGIT estimate of the standard error of $\gamma_{1o}$ was slightly larger (.2290) than the

VARCL estimate (0.1979). However, both their values were larger than the SPSS estimate (.1726) of S.E.($\gamma_{1o}$).

The estimates variance-covariance components of the random effects produced by MULTILOGIT are slightly larger than VARCL estimates (see Table 4-2b). In addition, the VARCL approach provided a test (by providing an estimate and its standard error) of the hypothesis of no variation across schools in: (a) $B_{oj}$, the adjusted school mean, $H_o{:}\tau_{oo} = 0$; (b) $B_{1j}$, the SES regression coefficient, $H_o{:}\tau_{11} = 0$; and (c) the covariance random effects between $B_{oj}$ and $B_{1j}$, $H_o{:}\tau_{o1} = 0$.

The t-statistic test of significance at p =.05 level implies no variation across the SES regression coefficients (t = 1.3152), and no significance covariance exists between the adjusted mean, $B_{oj}$, and the SES regression coefficients, $B_{1j}$ (t = -0.4518). The results of the test suggest that the variation across the adjusted mean is significant (t = 7.48886). This result is similar to that obtained in MLR model 1.

Thus, in analyzing this data, the regression coefficient of the SES slopes for the within-group model might well be fixed (i.e., changing the Equation 5.7 into $B_{1j} = \gamma_{1o}$). This decision cannot be made utilizing the MULTILOGIT program since it does not report any testing for the random effects portion of the model.

Table 4-2b.--Estimated variance-covariance components of the random effects and S.E. for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{00}}$, and $\tau_{01}$ (given in parentheses) for MLR model 2 using the VARCL and MULTILOGIT estimation methods.

| | Estimation Method | | | |
|---|---|---|---|---|
| | Quasi-likelihood (VARCL) | | Empirical Bayesian (MULTILOGIT) | |
| Intercept | 0.926374* (.128522) | | 1.01563 (NA) | |
| SES SLOPES | -0.10108+ (.223707) | 0.2114** (.34959) | .0572946 (NA) | .358306 (NA) |

Note
(i) NA-not given by the program.
(ii) *-significant at .05 ($H_0$: $\tau_{00}$=0, t-statistic=.96248/.1285=7.4889).
(iii) **-not significant at .05 ($H_0$: $\tau_{11}$=0, t-statistic=.4598/.3496=1.3152).
(iv) +-not significant at .05 ($H_0$:$\tau_{01}$=0, t-statistic=-.10108/.22371= -.452).

## The Results of MLR Model 3

The results of this analysis (Table 4-3a) show that the regression coefficient estimates of the within-school variable (i.e., intercept, $\gamma_{00}$, SES slope $\gamma_{10}$, and MSES slope $\gamma_{01}$) using the SPSS and the VARCL approach are close in values. The absolute MULTILOGIT estimate, however, is larger than both the SPSS and the VARCL estimates.

The figures in Table 4-3a illustrate that the standard error estimate for both the within- and the between-school variable regression coefficients (i.e., S.E.($\gamma_{00}$), S.E.($\gamma_{10}$), and S.E.($\gamma_{01}$)) for the VARCL and the MULTILOGIT approach are closer in value to each other. Their estimates of standard error were larger than the SPSS estimate.

The results of the variance-covariance components of the random effects estimates, their hypotheses tests, and the comparison between the different approaches for MLR model 3

Table 4-3a.--Estimated regression coefficient and standard error (given in parentheses) for MLR model 3 using different estimation methods.

| | Estimation Method | | |
|---|---|---|---|
| | M.L. (SPSS) | Quasi-likelihood (VARCL) | Empirical Bayesian (MULTILOGIT) |
| Intercept, $\gamma_{00}$ | -1.7875 | -1.795934 | -2.00627 |
| | (.0879) | (NA) | (.1655765) |
| SES, $\gamma_{10}$ | -.2454 | -0.294160 | -0.399041 |
| | (.2202) | (0.247753) | (.2440238) |
| MSES, $\gamma_{01}$ | -1.2960 | -1.315930 | -1.32994 |
| | (.3164) | (0.459423) | (0.4655073) |

Note: NA-not reported by the program.

are similar to the previous results for MLR model 1 and 2 (Table 4-3b).

## The Results of MLR Model 4

The estimates for the regression coefficients associated with interaction term, $\gamma_{11}$, for MULTILOGIT (.75059) and VARCL (.75448) are close in value (Table 4-4a). Both these estimates are larger than the SPSS (.6560) estimate.

Table 4-4a again shows that the MULTILOGIT and VARCL estimates of standard error of $\gamma_{11}$ are larger than the SPSS estimate.

The MULTILOGIT and VARCL estimates of $\tau_{00}$, and $\tau_{11}$ are close in value (Table 4-4b). However, again the MULTILOGIT estimates are larger than the VARCL estimates.

Table 4-3b.--Estimated variance-covariance components of the random effects and S.E. for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{00}}$, and $\tau_{01}$ (given in parentheses) for MLR model 3 using the VARCL and MULTILOGIT estimation methods.

| | Estimation Method | | | |
|---|---|---|---|---|
| | Quasi-likelihood (VARCL) | | Empirical Bayesian (MULTILOGIT) | |
| Intercept | 0.86048* (.129164) | | 0.924759 (NA) | |
| SES SLOPES | 0.17049+ (.256121) | 0.4314** (.35782) | 0.122416 (NA) | .319217 (NA) |

Note
(i) NA-not reported by the program.
(ii) *-significant at .05 ($H_0$: $\tau_{00}$=0, t-statistic=.92762/.1292 =7.1817).
(iii) **-not significant at .05 ($H_0$: $\tau_{11}$=0, t-statistic=.65683/.3578=1.836).
(iv) +-not significant at .05 ($H_0$: $\tau_{01}$=0, t-statistic=.17049/.25612=0.6657).

Table 4-4a.--Estimated regression coefficient and standard error (given in parentheses) for MLR model 4 using different estimation methods.

| | Estimation Method | | |
|---|---|---|---|
| | M.L. (SPSS) | Quasi-likelihood (VARCL) | Empirical Bayesian (MULTILOGIT) |
| Intercept, $\gamma_{00}$ | -1.7973 (.0875) | -1.767842 (NA) | -2.08383 (.1747097) |
| SES, $\gamma_{10}$ | -.4097 (.2383) | -0.387944 (0.244380) | -0.541050 (.2478449) |
| MSES, $\gamma_{01}$ | -1.1891 (.3254) | -1.152334 (0.456739) | -1.23976 (0.469006) |
| URB X SES, $\gamma_{11}$ | 0.6560 (.3568) | 0.754480 (0.405380) | 0.750593 (0.453603) |

Note: NA-not reported by the program.

Table 4-4b.--Estimated variance-covariance components of the random effects and S.E. for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{00}}$, and $\tau_{01}$ (given in parentheses) for MLR model 4 using the VARCL and MULTILOGIT estimation methods.

| | Estimation Method | | | |
|---|---|---|---|---|
| | Quasi-likelihood (VARCL) | | Empirical Bayesian (MULTILOGIT) | |
| Intercept | 0.92129* (.127039) | | 0.975108 (NA) | |
| SES SLOPES | 0.33158+ (.213225) | 0.1194** (.31258) | 0.215233 (NA) | .243533 (NA) |

Note
(i) NA-not reported by the program.
(ii) *-significant at .05 ($H_0$: $\tau_{00}$=0, t-statistic=.95984/.1271=7.5526).
(iii) **-not significant at .05 ($H_0$: $\tau_{11}$=0, t-statistic=.3455/.31268=1.105).
(iv) +-significant at .05 ($H_0$: $\tau_{01}$=0, t-statistic= .57583 / .21323 = 2.701).

## The Results of MLR Model 5

The results, shown in Table 4-5a, with respect to the $\gamma_{00}$ and $\gamma_{10}$ parameters are identical to the results of MLR model 3. However, comparing the regression coefficient estimates for the school-level variables (i.e. $\gamma_{01}$, $\gamma_{02}$, $\gamma_{03}$, $\gamma_{04}$ , $\gamma_{05}$, and $\gamma_{11}$), the results suggest that for $\gamma_{03}$, $\gamma_{04}$ , $\gamma_{05}$, and $\gamma_{11}$ the SPSS and VARCL estimates are close in value.

The smallest estimates are observed in the SPSS approach, while the largest estimates are observed in the MULTILOGIT approach. The VARCL estimates are close to SPSS estimates, and the estimates of the standard errors of both the within- and the between-school regression coefficient variables for MULTILOGIT and VARCL are also very close in value.

However,the MULTILOGIT estimates of the standard errors are consistently slightly larger than the VARCL estimates. For

Table 4-5a.--Estimated regression coefficient and standard error (given in parentheses) for MLR model 5 using different estimation methods.

| | Estimation Method | | | | | |
|---|---|---|---|---|---|---|
| | M.L. (SPSS) | t-STAT | Quasi-likelihood (VARCL) | t-STAT | Empirical Bayesian (MULTILOGIT) | t-STAT |
| Intercept, $\gamma_{oo}$ | -1.8357 | | -1.82266 | | -2.10500 | |
| | (.0906) | | (NA) | | (.1726824) | |
| SES, $\gamma_{1o}$ | -.3810 | -1.63 | -0.391991 | -1.62 | -0.536057 | -2.20 |
| | (.2340) | | (.24151) | | (.2437511) | |
| (URB/RRL), $\gamma_{o1}$ | -.4444 | -1.56 | -0.580533 | -1.23 | -0.423996 | -0.89 |
| | (.2844) | | (.473119) | | (.47761909) | |
| (CENTRAL), $\gamma_{o2}$ | .0919 | 0.36 | 0.073548 | 0.17 | 0.0415828 | 0.09 |
| | (.2557) | | (.425883) | | (.44582059) | |
| (NORTH), $\gamma_{o3}$ | .8605 | 4.14 | 0.867967 | 2.18 | 1.03390 | 2.46 |
| | (.2078) | | (.399045) | | (.42000833) | |
| (SOUTH), $\gamma_{o4}$ | .3835 | 1.61 | 0.382027 | 0.93 | 0.412540 | 0.94 |
| | (.2389) | | (.412146) | | (.43958617) | |
| MSES, $\gamma_{o5}$ | -.8753 | -2.39 | -0.902494 | -1.70 | -0.985701 | -1.84 |
| | (.3670) | | (.531539) | | (.53678115) | |
| URB X SES, $\gamma_{11}$ | .6794 | 1.86 | 0.700078 | 1.60 | 0.740221 | 1.63 |
| | (.3644) | | (.438400) | | (.453797) | |

Note: NA-not reported by the program.

example, the MULTILOGIT estimates of the standard errors for $\gamma_{1o}$, $\gamma_{o1}$, $\gamma_{o2}$, $\gamma_{o3}$, $\gamma_{o4}$ , $\gamma_{o5}$, and $\gamma_{11}$ are .244, .478, .446, .420, .440, .537, and .454 while the VARCL estimates are .242, .473, .426, .399, .412, .532, and .438, respectively. Both the VARCL and the MULTILOGIT estimates of the standard error are also much larger than the SPSS estimates.

The t-statistic computed for each of the regression coefficients, $\gamma$'s, for the three estimation methods is shown in Table 4-5a. This t-statistic provides a test of significance of the regression coefficient of the model. The null hypothesis is $H_o$: $\gamma_{ij} = 0$. This test helps resolve whether there is a significant relationship between the micro (or macro) covariate variables and the dependent variables. In this case, the objective of the t-statistic test is to compare the decisions made regarding the micro and macro covariate variables (i.e., either rejecting or accepting the null hypothesis) in the three methods of estimations.

The results of the covariate hypothesis testing for the micro variable show that the SPSS and VARCL estimation procedures each produced the same conclusion (except with respect to MSES variable). However, a different conclusion was reached using the MULTILOGIT estimation method. For example, the t-test of the SES regression coefficient variable, $\gamma_{1o}$, of the SPSS (t = -1.63) and the VARCL (t = -1.62) estimation procedures indicate that the null hypothesis cannot be rejected, while the MULTILOGIT (t = -2.2) method of estimation rejected the null hypothesis. Every hypothesis was tested at a .05 level of significance (see Table 4-5a).

The results of the covariate hypothesis tests for the macro variables show that the MULTILOGIT and VARCL estimation procedures reached the same conclusion unlike the SPSS estimation method. For example, at a .05 level of significance the t-test of MSES regression coefficient variable, $\gamma_{o5}$, indicates that both the MULTILOGIT (t = -1.84) and VARCL (t = -1.70) estimation procedures could not reject the null hypothesis, while the SPSS (t = -2.39) method of estimation rejected the null (see Table 4-5a).

The results of the variance-covariance components of the random effects estimates shown in Table 4-5b are similar to results of previous MLR models. As mentioned earlier, the

VARCL approach provides a variance-covariance components test of the random effects estimates. Again (as in MLR model 2) test results indicate no variation across groups in the SES within-school variable regression slopes (t = 1.1460). Thus, the SES slopes for the within-school model had to be fixed in order to compare the next two MLR models using the VARCL estimation method.

Table 4-5b.--Estimated variance-covariance components of the random effects and S.E. for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{00}}$, and $\tau_{01}$ (given in parentheses) for MLR model 5 using the VARCL and MULTILOGIT estimation methods.

| | Estimation Method | | | |
|---|---|---|---|---|
| | Quasi-likelihood (VARCL) | | Empirical Bayesian (MULTILOGIT) | |
| Intercept | 0.813675* (.125398) | | 0.914446 (NA) | |
| SES SLOPES | $0.353929^{+}$ (.211922) | 0.15396** (.34239) | 0.228719 (NA) | .214819 (NA) |

Note
(i) NA-not reported by the program.
(ii) *-significant at .05 ($H_0$: $\tau_{00}$=0, t-statistic=.902039/.125398=7.1934).
(iii) **-not significant at .05 ($H_0$: $\tau_{11}$=0, t-statistic=.39238/.3424=1.15).
(iv) +-significant at .05 ($H_0$: $\tau_{01}$=0, t-statistic=.59492/.21192=2.8073).

## The Results of Comparing Two MLRM Using VARCL Program

The two MLR models analyzed were denoted as model A and model B. Model B was a random intercept model, while model A had a combined random intercept and a random regression slope for the SES variable (This is similar to the VARCL model in Table 4-5a).

The comparison of these two models shows the effect of using one model over the other in decision making (i.e., rejecting or accepting the null hypotheses) regarding the effect of the student- or school-level variables on the dependent variable.

As can be seen from the results, the regression coefficients and the standard error (S.E. ($\gamma$'s)) estimates, $\gamma$'s, of model A and model B are different from each other. The t-statistic confirms this observation (see Table 4-6).

Table 4-6.--Estimated regression coefficient, and standard error (given in parentheses) for the **model A** having random intercept and random regression slope and **model B** having random intercept and fixed regression slope using VARCL estimation method.

| | MODEL A | T STATISTIC | MODEL B | T STATISTIC |
|---|---|---|---|---|
| Intercept, $\gamma_{00}$ | -1.82266<br>(NA) | | -1.812083<br>(NA) | |
| SES, $\gamma_{10}$ | -0.391991<br>(.241509) | -1.623 | -0.344946<br>(.232561) | -1.483 |
| (URB/RRL), $\gamma_{01}$ | -0.580533<br>(.473119) | -1.227 | -0.653884<br>(.444054) | -1.473 |
| (CENTRAL), $\gamma_{02}$ | 0.073548<br>(.425883) | 0.173 | 0.031822<br>(.413072) | 0.077 |
| (NORTH), $\gamma_{03}$ | 0.867967<br>(.399045) | 2.175 | 0.803747<br>(.412093) | 1.950 |
| (SOUTH), $\gamma_{04}$ | 0.382027<br>(.412146) | 0.927 | 0.245883<br>(.419792) | 0.586 |
| MSES, $\gamma_{05}$ | -0.902494<br>(0.531539) | -1.698 | -0.831582<br>(.495013) | -1.680 |
| URB X SES, $\gamma_{11}$ | 0.700078<br>(.438400) | 1.597 | 0.657214<br>(.399225) | 1.646 |

Note: NA-not reported by the program.

The t-statistic for $\gamma_{03}$ using model A suggests that $\gamma_{03}$ is significantly different from zero at a .05 significance level (t = 2.175), while for model B the test indicates that $\gamma_{03}$ is not significant ($t = 1.95$). Thus, the ability of the VARCL program to test the variance of the random effects is useful, whenever we want to account for the group membership effect.

## Results of the Simulated Data Analysis

First, the results of comparing the SPSS and VARCL estimation methods will be discussed with respect to: (a) the estimates of the macro parameters, $\gamma$'s ($\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$, $\gamma_{11}$), and (b) the estimates of the standard errors of the macro parameter.

Second, the effect of the following three simulated conditions will be evaluated: (a) number of units within each group, (b) the magnitude of the random regression slope variance in contrast to the intercept variance, and (c) the distribution of the random effects. For both the SPSS and VARCL estimation procedures, the above three conditions will be considered with respect to the following statistics: (a) the macro parameters estimates, and (b) the estimates of standard errors of the macro parameters.

Finally, the accuracy of the VARCL estimate of the variance-covariance components of the random effects and its estimate of the standard error will be discussed.

### Comparison of $\gamma$'s Between the SPSS and VARCL Estimation Methods

The purpose of this analysis was to compare the SPSS (standard single logistic regression model using Maximum-likelihood methods of estimation) and the VARCL (multilevel logistic regression model using Quasi-likelihood methods of estimation) properties of macro estimation.

Table 4-7 shows the true value, the mean of all four macro parameters and their standard errors of estimate, MSE of estimate, and the bias of both the SPSS and the VARCL estimation method. The statistical values of the macro parameters were obtained using 1200 replications. Similar statistical values were used to compare the properties and the accuracy of both estimation procedures for macro parameters under different experimental conditions (i.e., six cells) having 200 replications within each cell, presented in Tables 4-8 through 4-11.

The results of the analysis indicated that both the VARCL and the SPSS estimates of $\gamma$'s were statistically significantly biased at the significant level of $p = .05$. In fact, both estimation procedures underestimated the population parameters of $\gamma$'s. On the average, the estimates of $\gamma_{00}$ for both estimation methods were 13 percent smaller than the true value. Similarly, $\gamma_{01}$ was 12 percent smaller, $\gamma_{10}$ was 14 percent smaller and $\gamma_{11}$ was 12 percent smaller than their true values.

The VARCL and SPSS estimates of macro parameters were found, on the average, to be approximately equal for different statistics (i.e., mean, standard errors of estimate, MSE of estimate, and bias). A similar pattern of results was also detected with respect to estimates in the analysis of the real

Table 4-7.--The true macro parameter value, Mean, S.E[a]., MSE, and bias for estimated $\gamma$'s of the SPSS and VARCL estimation procedures[+].

| | | Estimation Method | |
|---|---|---|---|
| | | SPSS | VARCL |
| Macro Parameter | $\gamma_{00}$ | | |
| The True Value | | -1.800 | -1.800 |
| Mean of Estimate | | -1.569 | -1.572 |
| S.E. of Estimate | | .15 | .149 |
| MSE of Estimate | | .076 | .074 |
| Bias | | .231 | .228 |
| 95% CI Bias | | .22,.24 | .22,.24 |
| Percent of Bias | | 13% | 13% |
| Macro Parameter | $\gamma_{01}$ | | |
| The True Value | | -1.200 | -1.200 |
| Mean of Estimate | | -1.062 | -1.074 |
| S.E. of Estimate | | .163 | .163 |
| MSE of Estimate | | .045 | .043 |
| Bias | | .138 | .126 |
| 95% CI Bias | | .13,.15 | .12,.14 |
| Percent of Bias | | 12% | 11% |
| Macro Parameter | $\gamma_{10}$ | | |
| The True Value | | -0.500 | -0.500 |
| Mean of Estimate | | -.430 | -.428 |
| S.E. of Estimate | | .112 | .106 |
| MSE of Estimate | | .017 | .016 |
| Bias | | .070 | .072 |
| 95% CI Bias | | .06,.08 | .07,.08 |
| Percent of Bias | | 14% | 14% |
| Macro Parameter | $\gamma_{11}$ | | |
| The True Value | | .75 | .75 |
| Mean of Estimate | | .662 | .668 |
| S.E. of Estimate | | .144 | .140 |
| MSE of Estimate | | .028 | .026 |
| Bias | | -.088 | -.082 |
| 95% CI Bias | | -.10,-.08 | -.09,-.07 |
| Percent of Bias | | 12% | 11% |

Note
[a] Observed standard deviation of estimates
[+] From 1200 replications.

Table 4-8.--The true value, Mean, S.E[a]., MSE, and bias for estimated $\gamma_{\infty}$ by cell identification for the SPSS and the VARCL estimation procedure*.

Macro Parameter $\gamma_{\infty}$

| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -1.800 | -1.800 | -1.800 | -1.800 |
| Mean of Estimate | -1.555 | -1.557 | -1.586 | -1.590 |
| S.E. of Estimate | .175 | .174 | .169 | .170 |
| MSE of Estimate | .091 | .89 | .075 | .073 |
| Bias | .245 | .243 | .214 | .210 |
| 95% CI Bias | .22,.27 | .22,.27 | .19,.24 | .19,.23 |
| Percent of Bias | 14% | 14% | 12% | 12% |
| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -1.800 | -1.800 | -1.800 | -1.800 |
| Mean of Estimate | -1.556 | -1.560 | -1.546 | -1.553 |
| S.E. of Estimate | .141 | .137 | .118 | .121 |
| MSE of Estimate | .079 | .076 | .079 | .076 |
| Bias | .244 | .240 | .254 | .247 |
| 95% CI Bias | .22,.26 | .22,.26 | .24,.27 | .23,.27 |
| Percent of Bias | 14% | 13% | 14% | 14% |
| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -1.800 | -1.800 | -1.800 | -1.800 |
| Mean of Estimate | -1.587 | -1.589 | -1.582 | -1.585 |
| S.E. of Estimate | .155 | .156 | .127 | .125 |
| MSE of Estimate | .069 | .069 | .064 | .062 |
| Bias | .213 | .211 | .218 | .215 |
| 95% CI Bias | .19,.24 | .19,.23 | .20,.24 | .20,.23 |
| Percent of Bias | 12% | 12% | 12% | 12% |

Note
* -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{\infty}$=.995, $\tau_{11}$=.005).
RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{\infty}$=.85, $\tau_{11}$=.15).

Table 4-9.--The true value, Mean, S.E[a]., MSE, and bias for estimated $\gamma_{01}$ by cell identification for the SPSS and the VARCL estimation procedure[*].

| Macro Parameter | $\gamma_{01}$ | | | |
|---|---|---|---|---|
| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -1.200 | -1.200 | -1.200 | -1.200 |
| Mean of Estimate | -1.077 | -1.086 | -1.057 | -1.070 |
| S.E. of Estimate | .193 | .196 | .180 | .182 |
| MSE of Estimate | .052 | .051 | .053 | .050 |
| Bias | .123 | .114 | .143 | .130 |
| 95% CI Bias | .10,.15 | .09,.14 | .12,.17 | .11,.16 |
| Percent of Bias | 10% | 10% | 12% | 11% |
| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -1.200 | -1.200 | -1.200 | -1.200 |
| Mean of Estimate | -1.048 | -1.057 | -1.059 | -1.077 |
| S.E. of Estimate | .153 | .148 | .133 | .132 |
| MSE of Estimate | .046 | .042 | .037 | .032 |
| Bias | .152 | .143 | .141 | .123 |
| 95% CI Bias | .13,.17 | .12,.16 | .12,.16 | .11,.14 |
| Percent of Bias | 13% | 12% | 12% | 10% |
| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -1.200 | -1.200 | -1.200 | -1.200 |
| Mean of Estimate | -1.059 | -1.068 | -1.074 | -1.085 |
| S.E. of Estimate | .176 | .178 | .133 | .131 |
| MSE of Estimate | .051 | .049 | .034 | .030 |
| Bias | .141 | .132 | .126 | .115 |
| 95% CI Bias | .12,.17 | .11,.16 | .11,.14 | .10,.13 |
| Percent of Bias | 12% | 11% | 11% | 10% |

Note
[*] -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).
RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

Table 4-10.--The true value, Mean, S.E[a]., MSE, and bias for estimated $\gamma_{10}$ by cell identification for the SPSS and the VARCL estimation procedure*.

| Macro Parameter | $\gamma_{10}$ | | | |
|---|---|---|---|---|
| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -.50 | -.50 | -.50 | -.50 |
| Mean of Estimate | -.437 | -.437 | -.427 | -.424 |
| S.E. of Estimate | .133 | .131 | .138 | .135 |
| MSE of Estimate | .021 | .021 | .024 | .024 |
| Bias | .063 | .063 | .073 | .076 |
| 95% CI Bias | .05,.08 | .05,.08 | .05,.09 | .06,.10 |
| Percent of Bias | 13% | 13% | 15% | 15% |
| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -.50 | -.50 | -.50 | -.50 |
| Mean of Estimate | -.432 | -.433 | -.418 | -.413 |
| S.E. of Estimate | .080 | .052 | .081 | .078 |
| MSE of Estimate | .011 | .007 | .013 | .014 |
| Bias | .068 | .067 | .082 | .087 |
| 95% CI Bias | .06,.08 | .06,.08 | .07,.09 | .08,.10 |
| Percent of Bias | 14% | 13% | 16% | 17% |
| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | -.50 | -.50 | -.50 | -.50 |
| Mean of Estimate | -.436 | -.434 | -.429 | -.427 |
| S.E. of Estimate | .139 | .135 | .078 | .072 |
| MSE of Estimate | .023 | .023 | .011 | .011 |
| Bias | .064 | .066 | .071 | .073 |
| 95% CI Bias | .04,.08 | .05,.09 | .06,.08 | .06,.08 |
| Percent of Bias | 13% | 13% | 14% | 15% |

Note
* -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).
RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

Table 4-11.--The true value, Mean, S.E[a]., MSE, and bias for estimated $\gamma_{11}$ by cell identification for the SPSS and the VARCL estimation procedure*.

Macro Parameter $\gamma_{11}$

| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | .75 | .75 | .75 | .75 |
| Mean of Estimate | .677 | .677 | .654 | .661 |
| S.E. of Estimate | .172 | .167 | .178 | .184 |
| MSE of Estimate | .035 | .033 | .041 | .042 |
| Bias | -.073 | -.073 | -.096 | -.089 |
| 95% CI Bias | -.1,-.05 | -.1,-.05 | -.12,-.1 | -.11,-.1 |
| Percent of Bias | 10% | 10% | 13% | 12% |

| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | .75 | .75 | .75 | .75 |
| Mean of Estimate | .649 | .651 | .656 | .664 |
| S.E. of Estimate | .085 | .071 | .095 | .084 |
| MSE of Estimate | .017 | .015 | .018 | .014 |
| Bias | -.101 | -.099 | -.094 | -.086 |
| 95% CI Bias | -.1,-.09 | -.1,-.09 | -.11,-.1 | -.1,-.07 |
| Percent of Bias | 13% | 13% | 13% | 11% |

| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The True Value | .75 | .75 | .75 | .75 |
| Mean of Estimate | .674 | .683 | .662 | .669 |
| S.E. of Estimate | .195 | .193 | .089 | .082 |
| MSE of Estimate | .044 | .041 | .016 | .013 |
| Bias | -.076 | -.067 | -.088 | -.081 |
| 95% CI Bias | -.1,-.05 | -.1,-.04 | -.1,-.08 | -.1,-.07 |
| Percent of Bias | 10% | 9% | 12% | 11% |

Note
* -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).
RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

data.

In addition, the standard errors for estimated $\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$ were found to be quite close for both estimation procedures. However, the sampling distribution of these macro parameter estimates were different for both estimation methods (see Appendix 4-1 through 4-4).

A comparison of the MSE for both procedures indicated that the estimates of MSE for $\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$ were very close to zero for VARCL (.074, .043, .016 and .026) and slightly smaller than the SPSS estimates (.076, .045, .017 and .028). A similar conclusion may be deduced if the results of the macro parameters compared were under different experimental conditions, see Table 4-8 through 4-11.

The results in Table 4-8 through 4-11 implied that both the SPSS and VARCL programs estimates of four macro parameters were statistically significantly biased at the significant level of $p = .05$. The bias ranged between 9% and 17% smaller than the true value. Further investigation was carried out on the simulated program. First, all the commands of the simulated program and the transformation of the dependent variable into binary outcomes were rechecked. The sample size of the number of schools were increased and the programs were executed again. The VARCL estimates proved to be still significantly biased even with the increased sample size. The simulated program was then subdivided and each part analyzed separately. The distribution of the random effects was checked and found to be normal. The simulation was run for the fixed multi-level logistic regression model excluding the random effects ($U_{oj}$, $U_{1j}$) from the school level model. The findings indicated that the VARCL estimates of the $\gamma$s were unbiased. The variance of $U_{oj}$ and $U_{1j}$ were estimated as zero. These results indicated that the simulation program was working correctly with the fixed model.

Three models were then run:

(a) Model A (Random Intercept Logistic Regression Model) represented as

$$\alpha_{ij} = B_{oj}$$

$$B_{oj} = \gamma_{oo} + U_{oj} \quad U_{oj} \sim N(0, \tau_{oo}) \ ;$$

(b) Model B (Fixed Intercept Logistic Regression Model) represented as

$$\alpha_{ij} = B_{oj}$$

$$B_{oj} = \gamma_{oo} \ ; \text{ and}$$

(C) Model C (Random Effects Intercept Logistic Regression Model) represented as

$$\alpha_{ij} = B_{oj}$$

$$B_{oj} = U_{oj} \quad U_{oj} \sim N(0, \tau_{oo})$$

This was done in order to isolate the effect of the independent variables of the school ($Z_j$) and the students ($X_{ij}$) from the estimation parameters and help identify the source of the problem.

The results in Table 4-12 using Model B, the fixed model, showed that the VARCL estimates $\gamma_{oo}$ and $\tau_{oo}$ were very close to the true values. However, the results using the random intercept model, Model A, showed that the VARCL estimates,

Table 4-12.--Estimated $\gamma_{00}$ and estimated $\tau_{00}$ for different models using VARCL estimation methods*.

| Estimated Parameter | True Value | Estimated Value |
|---|---|---|
| | **MODEL A** | |
| $\gamma_{00}$ | -1.800 | -1.560 |
| Std Dev. | | .019 |
| Maximum | | -1.52 |
| Minimum | | -1.59 |
| $\tau_{00}$ | .85 | 0.710 |
| Std Dev. | | .048 |
| Maximum | | .78 |
| Minimum | | .64 |
| | **MODEL B** | |
| $\gamma_{00}$ | -1.800 | -1.789 |
| Std Dev. | | .032 |
| Maximum | | -1.74 |
| Minimum | | -1.84 |
| $\tau_{00}$ | 0.00 | 0.001 |
| Std Dev. | | .003 |
| Maximum | | .01 |
| Minimum | | .000 |
| | **MODEL C** | |
| $\gamma_{00}$ | 0.00 | -0.006 |
| Std Dev. | | .021 |
| Maximum | | .03 |
| Minimum | | -.04 |
| $\tau_{00}$ | .85 | 0.605 |
| Std Dev. | | .032 |
| Maximum | | .65 |
| Minimum | | .55 |

Note: * 10 replications were used for each model.

both $\gamma_{00}$ and $\tau_{00}$, were biased and under estimated the true values. These results were based on ten replications using 130 schools with 60 students within each school (N=7800) in each replication. $\tau_{00}$ was set to .85, $\gamma_{00}$ was set to -1.800 and $U_{oj}$ was centered (i.e., mean of zero) for these preliminary analyzes.

Similar analyzes using model A were also performed for both the SPSS and VARCL estimation methods, where $\gamma_{00}$ was set to -1.00 (rather than -1.800 in the earlier analysis) and $\tau_{00}$ was set to have different values: .04, 1.0, .30, .50, .70, .85. This was done because of the concern that earlier extreme magnitudes of the simulated value for $\tau_{00}$, and $\gamma_{00}$ may have caused the VARCL estimate to be biased and inconsistent.

The results in Table 4-13 confirmed earlier findings. In fact, the results also indicated that as the true value for $\tau_{00}$ increased from .04 to .85 the magnitude of bias for the VARCL estimate of both $\gamma_{00}$ and $\tau_{00}$ increased. Both estimates underestimated the true values. This result was based on 5 replications for each situation on a total of 7800 subjects in each replication ( j=130 groups, i=60 subjects). In fact, the real data analysis that was based on 59 schools consisting of 1244 students also showed that the VARCL estimates were of smaller magnitude than the MULTILOGIT estimates of Wong & Mason (see Table 4-1a and 4-1b).

In addition, the results in Table 4-13 also indicated that the SPSS estimates of $\gamma_{00}$ moved further away form the true value (i.e., the magnitude of bias increases) as the true

Table 4-13.--Estimated $\gamma_{00}$ and estimated $\tau_{00}$ using VARCL program*.

| Estimated Parameter | VARCL | SPSS | VARCL | SPSS | VARCL | SPSS |
|---|---|---|---|---|---|---|
| | Estimation Methods | | | | | |
| **Case** | **1** | | **2** | | **3** | |
| $\gamma_{00}$ | | | | | | |
| **True Value** | **-1.00** | **-1.00** | **-1.00** | **-1.00** | **-1.00** | **-1.00** |
| Estimate | -.986 | -.986 | -.966 | -.966 | -0.94 | -.94 |
| Stand. Dev. | .035 | .035 | .014 | .014 | 0.019 | .019 |
| Maximum | -.93 | -.93 | -.951 | -.951 | -0.917 | -.917 |
| Minimum | -1.04 | -1.04 | -.983 | -.983 | -0.964 | -.964 |
| $\tau_{00}$ | | | | | | |
| **True Value** | **.04** | **.04** | **.100** | **.100** | **.300** | **.300** |
| Estimate | .037 | NA | .099 | NA | 0.291 | NA |
| Stand. Dev. | .013 | NA | .026 | NA | 0.036 | NA |
| Maximum | .06 | NA | .133 | NA | 0.332 | NA |
| Minimum | .02 | NA | .063 | NA | 0.235 | NA |
| **Case** | **4** | | **5** | | **6** | |
| $\gamma_{00}$ | | | | | | |
| **True Value** | **-1.00** | **-1.00** | **-1.00** | **-1.00** | **-1.00** | **-1.00** |
| Estimate | -0.900 | -0.900 | -0.879 | -0.879 | -0.834 | -0.834 |
| Stand. Dev. | 0.018 | 0.018 | 0.029 | 0.029 | 0.016 | 0.016 |
| Maximum | -0.872 | -0.872 | -0.838 | -0.838 | -0.817 | -0.817 |
| Minimum | -0.919 | -0.919 | -0.915 | -0.915 | -0.854 | -0.854 |
| $\tau_{00}$ | | | | | | |
| **True Value** | **.500** | **.500** | **.700** | **.700** | **.85** | **.85** |
| Estimate | 0.420 | NA | 0.581 | NA | 0.609 | NA |
| Stand. Dev. | 0.030 | NA | 0.042 | NA | 0.014 | NA |
| Maximum | 0.459 | NA | 0.641 | NA | 0.619 | NA |
| Minimum | 0.374 | NA | 0.543 | NA | 0.585 | NA |

Note: * 5 replications were used for each case with the exception of case number 1 where 10 replications were used.

value for $\tau_{oo}$ increased from .04 to .85. The reason for this was because the model generated in the study did not fit the SPSS model. The SPSS program does not account for the effect of the random effects in the model. In fact, the SPSS estimates of $\gamma$s for the fully fixed model (generated by excluding the random effects, $U_{oj}$ and $U_{1j}$ from the full random effects model, based on j=60 and i=60) were unbiased. see Table 4-14.

The Fully Fixed model.

The with-group model is represented as

$$\alpha_{ij} = B_{oj} + B_{1j} X_{ij}$$

The between-group model is represented as

$$B_{oj} = \gamma_{oo} + \gamma_{o1} Z_j$$

$$B_{1j} = \gamma_{1o} + \gamma_{11} Z_j$$

Table 4-14.--The true value, estimated $\gamma$'s, for three replications having the fully fixed model using the SPSS estimation method.

| Estimated Parameter | True Value | Replication First repl. | Second repl. | Third repl. | Mean of Estimates |
|---|---|---|---|---|---|
| $\gamma oo$ | -1.800 | -1.738285 | -1.893364 | -1.845619 | -1.826 |
| $\gamma 1o$ | -.50 | -0.416737 | -0.627042 | -0.401306 | -0.482 |
| $\gamma 11$ | .75 | 0.789271 | 0.762302 | 0.754291 | 0.769 |
| $\gamma o1$ | -1.20 | -1.156474 | -1.291167 | -1.189164 | -1.212 |

The results of the above analyses indicated that perhaps the VARCL program is actually biased and inconsistent in estimating $\gamma_{00}$ and $\tau_{00}$ for the random effects model. Communication with Longford (1992) confirmed that a bias existed in the estimation of $\gamma_{00}$ by the VARCL program. The negative bias of the ML estimator of $\tau_{00}$ was partly due to the approximation of the maximum likelihood method. Therefore, the results of the analysis on real and simulated data using VARCL program should be looked at with caution as the program gives a negative bias estimator of both $\gamma_{00}$ and $\tau_{00}$.

Comparison of the Standard Error of the $\gamma$'s Between the SPSS and VARCL Estimation Methods

One of the aims of the study was to compare estimates of the standard errors for the macro parameter, $\gamma$'s (i.e., $\gamma_{01}$, $\gamma_{10}$, and $\gamma_{11}$), using the SPSS and the VARCL estimation methods. These three macro parameters were used in this analysis because the VARCL program printout did not report the standard error of the intercept coefficient, $\gamma_{00}$.

The estimates of the standard error of macro parameters were obtained for both estimation methods. The mean, the standard error, the MSE, and the bias of estimated standard errors of the three macro parameters (i.e., $\gamma_{01}$, $\gamma_{10}$, and $\gamma_{11}$) for each estimation method are shown in Table 4-15. In addition, its true value was obtained from the observed standard error of the estimated macro parameters from the 1200 replications shown in Table 4-7. Similar statistical values were used to compare the properties and the accuracy of both estimation procedures for the estimated standard error of $\gamma_{01}$, $\gamma_{10}$, and $\gamma_{11}$ under six experimental conditions (see Tables 4-16 through 4-18). The observed standard error of the three estimated macro parameters in Tables 4-9 through 4-11 were used as the true value of the standard error in the Tables 4-16 through 4-18.

The results showed that, on the average, the estimates of the standard error of $\gamma$'s for VARCL were consistently larger than the SPSS estimates of the standard error.

In addition, the results in Table 4-15 showed, on the average, that the standard deviation for the estimates of the standard errors of $\gamma$'s for VARCL and SPSS were close. The results also indicated that the VARCL estimates of the standard error for $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$ were less biased than the SPSS estimates. In fact, the 95% confidence interval for bias, shown in Tables 4-16 through 4-18, indicated generally that the VARCL estimate of the standard errors of the three macro parameters was unbiased. Furthermore, the MSE for the estimated standard error for both methods of $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$ were also found to be quite close. However, the MSE for the estimated standard error of $\gamma_{01}$ was slightly smaller for VARCL than SPSS for each of the different experimental conditions, see Table 4-16.

The probability of type I error was also investigated. The Z-score ($Z = (\hat{\gamma} - \gamma_t)/S.E.(\hat{\gamma})$) was calculated for each replication, and the probability of the type I error rate

Table 4-15.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error for macro parameters by the SPSS and VARCL estimation procedures+.

| | Estimation Method | |
|---|---|---|
| | SPSS | VARCL |
| Macro parameter $\gamma_{01}$ | | |
| The True standard error | .163 | .163 |
| Mean of Estimate | .101 | .160 |
| S.E. of Estimate | .044 | .033 |
| MSE of Estimate | .006 | .001 |
| Bias | -.062 | -.003 |
| 95% CI Bias | -.063,-.060 | -.005,-.001 |
| Percent of Bias | 38% | 2% |
| Macro parameter $\gamma_{10}$ | | |
| The True standard error | .112 | .106 |
| Mean of Estimate | .090 | .103 |
| S.E. of Estimate | .041 | .040 |
| MSE of Estimate | .002 | .002 |
| Bias | -.022 | -.003 |
| 95% CI Bias | -.024,-.020 | -.005,-.001 |
| Percent of Bias | 20% | 3% |
| Macro parameter $\gamma_{11}$ | | |
| The True standard error | .144 | .140 |
| Mean of Estimate | .112 | .126 |
| S.E. of Estimate | .051 | .051 |
| MSE of Estimate | .004 | .003 |
| Bias | -.032 | -.014 |
| 95% CI Bias | -.034,-.030 | -.016,-.012 |
| Percent of Bias | 22% | 10% |

Note
+ 1200 replications were performed in each method.
* The true value was obtained from the observed standard error of the estimated macro parameters by VARCL and SPSS estimation methods from the 1200 replications shown in Table 4-7.

Table 4-16.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error of $\gamma_{01}$ by cell identification for the SPSS and the VARCL estimation procedure**.

| Macro Parameter | $\gamma_{01}$ | | | |
|---|---|---|---|---|
| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .193 | .196 | .180 | .182 |
| Mean of Estimate | .144 | .193 | .143 | .186 |
| S.E. of Estimate | .012 | .020 | .011 | .019 |
| MSE of Estimate | .003 | .000 | .001 | .000 |
| Bias | -.049 | -.003 | -.037 | .004 |
| 95% CI Bias | -.05,-.04 | -.01,0.00 | -.04,-.035 | 0.00,.01 |
| Percent of Bias | 25% | 2% | 21% | 2% |
| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .153 | .148 | .133 | .132 |
| Mean of Estimate | .058 | .139 | .058 | .130 |
| S.E. of Estimate | .003 | .013 | .003 | .012 |
| MSE of Estimate | .009 | .000 | .006 | .000 |
| Bias | -.095 | -.009 | -.075 | -.002 |
| 95% CI Bias | -.1,-.1 | -.01,0.00 | -.08,-.08 | -.004,0.0 |
| Percent of Bias | 62% | 6% | 56% | 2% |
| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .176 | .178 | .133 | .131 |
| Mean of Estimate | .143 | .186 | .058 | .128 |
| S.E. of Estimate | .011 | .020 | .003 | .017 |
| MSE of Estimate | .001 | .000 | .006 | .000 |
| Bias | -.033 | .008 | -.075 | -.003 |
| 95% CI Bias | -.031,-.03 | 0.0,.01 | -.08,-.08 | -.01,0.0 |
| Percent of Bias | 19% | 5% | 56% | 2% |

Note

* The true standard error value was obtained from the observed standard error of the estimated $\gamma_{01}$ macro parameter by cell identification for the SPSS and the VARCL estimation procedure shown in Table 4-9.

** -200 replications were performed within each cell.

ND -normal distribution of the random effects.

TD -t-distribution of the random effects.

n10 -10 subjects within each group.

n60 -60 subjects within each group.

RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).

RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

Table 4-17.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error of $\gamma_{10}$ by cell identification for the SPSS and the VARCL estimation procedure**.

| Macro Parameter | $\gamma_{10}$ | | | |
|---|---|---|---|---|
| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .133 | .131 | .138 | .135 |
| Mean of Estimate | .130 | .138 | .131 | .143 |
| S.E. of Estimate | .008 | .011 | .009 | .014 |
| MSE of Estimate | .000 | .000 | .000 | .000 |
| Bias | -.003 | .007 | -.007 | .008 |
| 95% CI Bias | -.01,0.0 | 0.0,.01 | -.01,0.0 | 0.0,.01 |
| Percent of Bias | 2% | 5% | 5% | 6% |
| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .080 | .052 | .081 | .078 |
| Mean of Estimate | .050 | .054 | .050 | .070 |
| S.E. of Estimate | .002 | .004 | .002 | .008 |
| MSE of Estimate | .001 | .000 | .001 | .000 |
| Bias | -.030 | .002 | -.031 | -.008 |
| 95% CI Bias | -.03,-.03 | .002,.002 | -.03,-.03 | -.01,0.0 |
| Percent of Bias | 4% | 4% | 38% | 10% |
| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .139 | .135 | .078 | .072 |
| Mean of Estimate | .131 | .143 | .050 | .069 |
| S.E. of Estimate | .008 | .013 | .002 | .008 |
| MSE of Estimate | .000 | .000 | .001 | .000 |
| Bias | -.008 | .008 | -.028 | -.003 |
| 95% CI Bias | -.01,0.0 | 0.0,.01 | -.03,-.03 | -.01,0.0 |
| Percent of Bias | 6% | 6% | 36% | 4% |

Note
* The true standard error value was obtained from the observed standard error of the estimated $\gamma_{10}$ macro parameter by cell identification for the SPSS and the VARCL estimation procedure shown in Table 4-10.
** -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).
RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

Table 4-18.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error of $\gamma_{11}$ by cell identification for the SPSS and the VARCL estimation procedure**.

| Macro Parameter | $\gamma_{11}$ | | | |
|---|---|---|---|---|
| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .172 | .167 | .178 | .184 |
| Mean of Estimate | .163 | .173 | .162 | .176 |
| S.E. of Estimate | .016 | .018 | .014 | .019 |
| MSE of Estimate | .000 | .000 | .000 | .000 |
| Bias | -.009 | .006 | -.016 | -.008 |
| 95% CI Bias | -.01,0.0 | 0.0,.01 | -.02,-.01 | -.01,0.0 |
| Percent of Bias | 5% | 4% | 9% | 4% |
| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .085 | .071 | .095 | .084 |
| Mean of Estimate | .062 | .068 | .061 | .083 |
| S.E. of Estimate | .003 | .005 | .003 | .008 |
| MSE of Estimate | .001 | .000 | .001 | .000 |
| Bias | -.023 | -.003 | -.034 | -.001 |
| 95% CI Bias | -.02,-.02 | -.003,-.003 | -.03,-.03 | -.002,0.0 |
| Percent of Bias | 27% | 4% | 36% | 1% |
| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| The true standard error | .195 | .193 | .089 | .082 |
| Mean of Estimate | .163 | .177 | .062 | .082 |
| S.E. of Estimate | .015 | .018 | .003 | .008 |
| MSE of Estimate | .001 | .001 | .001 | .000 |
| Bias | -.032 | -.016 | -.027 | .000 |
| 95% CI Bias | -.034,-.03 | -.02,-.01 | -.03,-.03 | 0.0,0.0 |
| Percent of Bias | 16% | 8% | 30% | 0% |

Note

* The true standard error value was obtained from the observed standard error of the estimated $\gamma_{11}$ macro parameter by cell identification for the SPSS and the VARCL estimation procedure shown in Table 4-11.

** -200 replications were performed within each cell.

ND -normal distribution of the random effects.

TD -t-distribution of the random effects.

n10 -10 subjects within each group.

n60 -60 subjects within each group.

RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).

RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

under a true null hypothesis ($H_o$: $\gamma = \gamma_t$ ) was determined by counting the frequency with which the Z-score exceeded the critical value for .05 significance level and dividing by the total number of replications (i.e. 1200).

The results in Table 4-19 showed that the probability of type I error rate under a true null for the VARCL tests of the macro parameters $\gamma$'s were, on the average, relatively smaller than the SPSS error rates.

However, both estimation methods gave unacceptable high type I error rates (i.e., $p > .05$). This was confirmed by further investigation of the probability of type I error rates under a true null hypothesis under the different experimental conditions (i.e., six cells), see Table 4-20.

Table 4-19.--The probability of type I error rates* for tests of macro estimators under a true null by the SPSS and VARCL estimation procedures+.

| Macro Parameter | Estimation Method | |
|---|---|---|
| | SPSS | VARCL |
| $\gamma_{01}$ | .438 | .136 |
| $\gamma_{10}$ | .223 | .146 |
| $\gamma_{11}$ | .261 | .165 |

Note
+ From 1200 replications were performed for each method.
* .05 significance level.

Similar results were also reported in several other research studies that compared the single-level regression

Table 4-20.--The probability of type I error rates[+] for tests of macro estimators under a true null by cell identification for the SPSS and the VARCL estimation procedure[•].

| Cell identification | (ND,n10,RRSS) | | (ND,n10,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| $\gamma_{01}$ | .255 | .075 | .275 | .125 |
| $\gamma_{10}$ | .085 | .065 | .110 | .090 |
| $\gamma_{11}$ | .125 | .080 | .125 | .105 |

| Cell identification | (ND,n60,RRSS) | | (ND,n60,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| $\gamma_{01}$ | .650 | .200 | .610 | .165 |
| $\gamma_{10}$ | .280 | .220 | .405 | .245 |
| $\gamma_{11}$ | .410 | .355 | .370 | .155 |

| Cell identification | (TD,n10,RRSL) | | (TD,n60,RRSL) | |
|---|---|---|---|---|
| Estimation Method | SPSS | VARCL | SPSS | VARCL |
| $\gamma_{01}$ | .225 | .090 | .610 | .160 |
| $\gamma_{10}$ | .075 | .065 | .385 | .190 |
| $\gamma_{11}$ | .155 | .110 | .380 | .185 |

Note
[+] -.05 significance level.
[•] -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.995, $\tau_{11}$=.005).
RRSL-large magnitude of the random regression slope variance to the intercept variance (i.e, $\tau_{00}$=.85, $\tau_{11}$=.15).

model with the multilevel regression model (Walsh, 1947; Aitkin, Anderson, and Hinde, 1981; Raudenbush and Bryk, 1989). These studies concluded that using the single level model instead of the multilevel model increased the probability of a type I error. In the case of the SPSS estimation procedure, this was due to the liberal t-statistic values caused by small standard error estimates for the regression coefficients when assuming the single-level logistic regression model.

The Effect of n on $\gamma$'s

The effect of the number of units within each group (i.e., n=10 versus n=60) for both estimation procedures was evaluated based on the following statistics: (a) the macro parameters estimates, and (b) the estimates of standard errors of the macro parameters.

The 1200 replications were split into two categories based on the number of units (subjects) within each group. This resulted in 600 replications in the first category where n=10, and 600 replications in the second category where n=60. In addition, Tables 4-8 through 4-11, show the effects of different sample sizes on the four macro parameters under different experimental conditions.

The effect of the number of units (subjects) within each group for both the VARCL and SPSS estimation methods on the macro parameter estimates was similar. As such, the following discussion of the effect of number of units on the macro parameters applies equally to both the VARCL and SPSS estimation procedures.

Examination of the bias estimates for $\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$, on the average, indicated a sightly negative effect of an increasing number of units within each group on the above macro parameters, see Table 4-21. The VARCL bias of $\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$ (.234, .127, .076 and -.089) when n=60 were larger (.221, .125, .068 and -.076) when n=10.

However, the standard error and MSE of the $\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$, and $\gamma_{11}$ estimates were smaller when n=60 than when n=10. For example, with n=10, and 60 the standard error and MSE for VARCL $\gamma_{10}$ macro parameter estimate dropped from .134 and .023 to .069 and .010, respectively.

## The Effect of n on the Estimated Standard Error of $\gamma$

The results of the analyze are shown in Tables 4-22 and 4-16 through 4-18. On the average, the estimated standard error for the three macro parameters was found to be smaller when n=60 than when n=10, for both estimation procedures.

In addition, the results also confirmed the earlier finding that, on the average, the estimated standard error of the $\gamma$'s of VARCL was consistently larger than the SPSS estimates.

The results in Tables 4-16 through 4-18 and 4-22 indicated that generally, VARCL estimate of the standard error of macro parameters were unbiased.

In addition, the results in Table 4-22 indicated that, on the average, increasing the number of subjects within each

Table 4-21.--The true value, Mean, S.E., MSE, and bias for estimated $\gamma$'s by the number of subject within each group for the SPSS and VARCL estimation procedure.

| | | Estimation Method | | | |
|---|---|---|---|---|---|
| | | SPSS | VARCL | SPSS | VARCL |
| Number Of Subject Within Each Group+ | | 10 | 10 | 60 | 60 |
| Macro Parameter | $\gamma_{00}$ | | | | |
| The True Value | | -1.800 | -1.800 | -1.800 | -1.800 |
| Mean of Estimate | | -1.576 | -1.579 | -1.561 | -1.566 |
| S.E. of Estimate | | .167 | .167 | .130 | .128 |
| MSE of Estimate | | .078 | .077 | .074 | .071 |
| Bias | | .224 | .221 | .239 | .234 |
| 95% CI Bias | | .21,.24 | .21,.23 | .23,.25 | .22,.24 |
| Percent of Bias | | 12% | 12% | 13% | 13% |
| Macro Parameter | $\gamma_{01}$ | | | | |
| The True Value | | -1.200 | -1.200 | -1.200 | -1.200 |
| Mean of Estimate | | -1.064 | -1.075 | -1.060 | -1.073 |
| S.E. of Estimate | | .183 | .185 | .140 | .137 |
| MSE of Estimate | | .052 | .050 | .039 | .035 |
| Bias | | .136 | .125 | .140 | .127 |
| 95% CI Bias | | .12,.15 | .11,.14 | .13,.15 | .12,.14 |
| Percent of Bias | | 11% | 10% | 12% | 11% |
| Macro Parameter | $\gamma_{10}$ | | | | |
| The True Value | | -.500 | -.500 | -.500 | -.500 |
| Mean of Estimate | | -.434 | -.432 | -.426 | -.424 |
| S.E. of Estimate | | .136 | .134 | .080 | .069 |
| MSE of Estimate | | .023 | .023 | .012 | .010 |
| Bias | | .066 | .068 | .074 | .076 |
| 95% CI Bias | | .05,.08 | .06,.08 | .07,.08 | .07,.08 |
| Percent of Bias | | 13% | 14% | 15% | 15% |
| Macro Parameter | $\gamma_{11}$ | | | | |
| The True Value | | .75 | .75 | .75 | .75 |
| Mean of Estimate | | .668 | .674 | .656 | .661 |
| S.E. of Estimate | | .182 | .181 | .090 | .079 |
| MSE of Estimate | | .040 | .039 | .017 | .014 |
| Bias | | -.082 | -.076 | -.094 | -.089 |
| 95% CI Bias | | -.10,-.07 | -.09,-.06 | -.10,-.09 | -.09,-.08 |
| Percent of Bias | | 11% | 10% | 13% | 12% |

Note: + 600 replications were performed for each estimation method.

Table 4-22.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error for macro parameters by the number of subject within each group for the SPSS and VARCL estimation procedures.

| | Estimation Method | | | |
|---|---|---|---|---|
| | SPSS | VARCL | SPSS | VARCL |
| Number Of Subject Within Each Group+ | 10 | 10 | 60 | 60 |
| Macro Parameter $\gamma_{01}$ | | | | |
| The True standard error | .183 | .185 | .140 | .137 |
| Mean of Estimate | .143 | .188 | .058 | .132 |
| S.E. of Estimate | .012 | .020 | .003 | .015 |
| MSE of Estimate | .002 | .000 | .007 | .000 |
| Bias | -.040 | .003 | -.082 | -.005 |
| 95% CI Bias | -.04,-.04 | .001,.005 | -.082,-.082 | -.007,-.003 |
| Percent of Bias | 22% | 2% | 59% | 4% |
| Macro Parameter $\gamma_{10}$ | | | | |
| The True standard error | .136 | .134 | .080 | .069 |
| Mean of Estimate | .131 | .141 | .050 | .064 |
| S.E. of Estimate | .008 | .013 | .002 | .010 |
| MSE of Estimate | .000 | .000 | .001 | .000 |
| Bias | -.005 | .007 | -.030 | -.005 |
| 95% CI Bias | -.005,-.005 | .005,.009 | -.03,-.03 | -.005,-.005 |
| Percent of Bias | 4% | 5% | 38% | 7% |
| Macro Parameter $\gamma_{11}$ | | | | |
| The True standard error | .182 | .181 | .090 | .079 |
| Mean of Estimate | .162 | .175 | .062 | .077 |
| S.E. of Estimate | .015 | .018 | .003 | .010 |
| MSE of Estimate | .001 | .000 | .001 | .000 |
| Bias | -.020 | -.006 | -.028 | -.002 |
| 95% CI Bias | -.022,-.018 | -.008,-.004 | -.028,-.028 | -.002,-.002 |
| Percent of Bias | 11% | 3% | 31% | 3% |

Note
+ 600 replications were performed for each estimation method.
* The true value was obtained from the observed standard error of the estimated macro parameters by the number of subject within each group for the SPSS and VARCL estimation procedures shown in Table 4-21.

group had slight effect on the standard deviation and MSE of the VARCL estimated standard error for $\gamma_{o1}$, $\gamma_{1o}$ and $\gamma_{11}$.

In addition, the results indicated that the SPSS estimated standard error of the macro parameters were statistically significantly biased. In fact, the absolute magnitude of bias increased as the sample size increased. For example, Table 4-22 indicated that, on the average, the percent of bias for $\gamma_{o1}$, $\gamma_{1o}$ and $\gamma_{11}$ when n=10 were 22%, 4%, and 11% respectively. This increased to 59%, 38%, and 31% respectively when n=60.

Finally, the results also indicated that for different sample sizes, the MSE for the standard error of $\gamma_{1o}$ and $\gamma_{11}$ for both estimation methods were very close. However, for $\gamma_{o1}$, the MSE of VARCL is smaller than the SPSS estimate.

## The Effect of the Random Effects Distribution on $\gamma$'s

The effect of having a normal distribution (ND) versus t-distribution (TD) of the random effects, $U_{oj}$ and $U_{1j}$, for both estimation procedures was evaluated on the basis of the following statistics: (a) the macro parameters estimates, and (b) the estimates of standard errors of the macro parameters.

The 1200 replications were again split into two categories based on distribution of the random effects (i.e., ND vs. TD). This resulted in 800 replications in the first category where the distribution of the random effects was normally distributed, and 400 replications in the second category where distribution of the random effects was t-distributed. This uneven balance of the replications between the two categories was caused by simulating a t-distribution of the random effects only under large values of the random regression slope variance (i.e., RRSL).

Using the replications within each category, the true value, mean, standard error, MSE, and bias of estimated macro parameters for each estimation method are shown in Table 4-23. In addition, Tables 4-8 through 4-11 show the effects of the random effects distribution on four macro parameters by different experimental conditions.

The results indicated that there were no clear effects of the random effects distribution on estimation of three macro parameters, $\gamma_{00}$, $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$.

## The Effect of the Random Effects Distribution on the Estimated Standard Error of $\gamma$'s

The averaged standard deviation, the MSE, and the bias of estimated standard error of the three macro parameters for both the estimation methods were compared. This was done in terms of having normal distribution (ND) versus t-distribution (TD). The results are shown in Tables 4-24 and 4-16 through 4-18.

The results indicated that having a normal distribution or a t-distribution of the random effects had no clear effect on the estimated standard error of the macro parameters.

Table 4-23.--The true value, Mean, S.E., MSE, and bias for estimated $\gamma$'s by the distribution of the random effects for the SPSS and VARCL estimation procedure.

| | Estimation Method | | | |
|---|---|---|---|---|
| | SPSS | VARCL | SPSS | VARCL |
| The distribution of the random effects | normal distribution[+] | | t-distribution[++] | |
| Macro Parameter | $\gamma_{00}$ | | | |
| The True Value | -1.800 | -1.800 | -1.800 | -1.800 |
| Mean of Estimate | -1.561 | -1.565 | -1.584 | -1.587 |
| S.E. of Estimate | .153 | .153 | .142 | .141 |
| MSE of Estimate | .081 | .078 | .067 | .065 |
| Bias | .239 | .235 | .216 | .213 |
| 95% CI Bias | .23,.25 | .23,.24 | .20,.23 | .20,.23 |
| Percent of Bias | 13% | 13% | 12% | 12% |
| Macro Parameter | $\gamma_{01}$ | | | |
| The True Value | -1.200 | -1.200 | -1.200 | -1.200 |
| Mean of Estimate | -1.060 | -1.073 | -1.066 | -1.076 |
| S.E. of Estimate | .166 | .167 | .156 | .156 |
| MSE of Estimate | .047 | .044 | .042 | .040 |
| Bias | .140 | .127 | .134 | .124 |
| 95% CI Bias | .13,.15 | .12,.14 | .12,.15 | .11,.14 |
| Percent of Bias | 12% | 11% | 11% | 10% |
| Macro Parameter | $\gamma_{10}$ | | | |
| The True Value | -.500 | -.500 | -.500 | -.500 |
| Mean of Estimate | -.429 | -.427 | -.432 | -.431 |
| S.E. of Estimate | .111 | .105 | .112 | .108 |
| MSE of Estimate | .017 | .016 | .017 | .017 |
| Bias | .071 | .073 | .068 | .069 |
| 95% CI Bias | .06,.08 | .07,.08 | .06,.08 | .06,.08 |
| Percent of Bias | 14% | 15% | 14% | 14% |
| Macro Parameter | $\gamma_{11}$ | | | |
| The True Value | .75 | .75 | .75 | .75 |
| Mean of Estimate | .659 | .663 | .668 | .676 |
| S.E. of Estimate | .139 | .136 | .152 | .148 |
| MSE of Estimate | .028 | .026 | .030 | .027 |
| Bias | -.091 | -.087 | -.082 | -.074 |
| 95% CI Bias | -.10,-.08 | -.10,-.08 | -.10,-.07 | -.09,-.06 |
| Percent of Bias | 12% | 12% | 11% | 10% |

Note
[+] 800 replications were performed for each estimation method.
[++] 400 replications were performed for each estimation method.

Table 4-24.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error for macro parameters by the distribution of the random effects for the SPSS and VARCL estimation procedure.

| | Estimation Method | | | |
|---|---|---|---|---|
| | SPSS | VARCL | SPSS | VARCL |
| The distribution of the random effects | normal distribution+ | | t-distribution++ | |
| Macro Parameter $\gamma_{01}$ | | | | |
| The True standard error | .166 | .167 | .156 | .156 |
| Mean of Estimate | .100 | .162 | .101 | .157 |
| S.E. of Estimate | .044 | .032 | .043 | .035 |
| MSE of Estimate | .006 | .001 | .005 | .001 |
| Bias | -.066 | -.005 | -.055 | .001 |
| 95% CI Bias | -.07,-.06 | -.007,-.003 | -.06,-.05 | -.003,.005 |
| Percent of Bias | 40% | 3% | 35% | 1% |
| Macro Parameter $\gamma_{10}$ | | | | |
| The True standard error | .111 | .105 | .112 | .108 |
| Mean of Estimate | .090 | .101 | .091 | .106 |
| S.E. of Estimate | .041 | .041 | .041 | .039 |
| MSE of Estimate | .002 | .002 | .002 | .001 |
| Bias | -.021 | -.004 | -.021 | -.002 |
| 95% CI Bias | -.023,-.02 | -.006,-.002 | -.025,-.017 | -.006,.002 |
| Percent of Bias | 19% | 4% | 19% | 2% |
| Macro Parameter $\gamma_{11}$ | | | | |
| The True standard error | .139 | .136 | .152 | .148 |
| Mean of Estimate | .112 | .125 | .113 | .130 |
| S.E. of Estimate | .052 | .052 | .051 | .050 |
| MSE of Estimate | .003 | .003 | .004 | .003 |
| Bias | -.027 | -.011 | -.039 | -.018 |
| 95% CI Bias | -.03,-.023 | -.015,-.007 | -.045,-.033 | -.022,-.014 |
| Percent of Bias | 19% | 8% | 26% | 12% |

Note
+ 800 replications were performed for each estimation method.
++ 400 replications were performed for each estimation method.
* The true value was obtained from the observed standard error of the estimated macro parameters by the distribution of the random effects for the SPSS and VARCL estimation procedures shown in Table 4-23.

The Effect of the RRS Variance Magnitude on $\gamma$'s

The effects of the magnitude of the random regression slope (RRS) variance to the intercept variance (RRSS versus RRSL) were evaluated for both estimation procedures on the basis of the following statistics: (a) the macro parameters estimates, and (b) the estimates of the standard errors of the macro parameters.

The 1200 replications were split into two categories based on magnitude of the random regression slope variance (i.e., RRSS vs. RRSL). This resulted in 400 replications in the first category where the random regression slope variance was small and 800 replications in the second category where the random regression slope variance was large. As mentioned earlier, this uneven balance of replications between the two categories was caused by simulating a t-distribution of the random effects only under a large random regression slope variance (i.e., RRSL), see Table 4-25.

In addition, Tables 4-8 through 4-11 show the effect of RRSS vs. RRSL on four macro parameters under different experimental conditions with 200 replications performed within each cell.

The results of the analyzes shown in Tables 4-8 through 4-11 indicated no clear effect of RRSS vs RRSL on the macro parameters, estimates $\gamma_{00}$ and $\gamma_{01}$. However, the results in Tables 4-10 and 4-11 indicated that the macro parameters estimates, $\gamma_{10}$ and $\gamma_{11}$ generally had a smaller MSE and bias under RRSS than RRSL. For example, when n=60 with RRSL, the

Table 4-25.--The true value, Mean, S.E., MSE, and bias for estimated $\gamma$'s by the magnitude of random regression slope variance for the SPSS and VARCL estimation procedure.

| | | Estimation Method | | | |
|---|---|---|---|---|---|
| | | SPSS | VARCL | SPSS | VARCL |
| The Magnitude of Random Regression Slope Variance to the Intercept Variance | | SMALL+ | SMALL+ | LARGE++ | LARGE++ |
| Macro Parameter | $\gamma_{00}$ | | | | |
| The True Value | | -1.800 | -1.800 | -1.800 | -1.800 |
| Mean of Estimate | | -1.555 | -1.559 | -1.575 | -1.579 |
| S.E. of Estimate | | .159 | .156 | .145 | .145 |
| MSE of Estimate | | .085 | .083 | .072 | .070 |
| Bias | | .245 | .241 | .225 | .221 |
| 95% CI Bias | | .23,.26 | .23,.26 | .22,.23 | .21,.23 |
| Percent of Bias | | 14% | 13% | 13% | 12% |
| Macro Parameter | $\gamma_{01}$ | | | | |
| The True Value | | -1.200 | -1.200 | -1.200 | -1.200 |
| Mean of Estimate | | -1.063 | -1.072 | -1.062 | -1.075 |
| S.E. of Estimate | | .174 | .174 | .157 | .157 |
| MSE of Estimate | | .049 | .047 | .044 | .040 |
| Bias | | .137 | .128 | .138 | .125 |
| 95% CI Bias | | .12,.15 | .11,.15 | .13,.15 | .11,.14 |
| Percent of Bias | | 11% | 11% | 12% | 10% |
| Macro Parameter | $\gamma_{10}$ | | | | |
| The True Value | | -.500 | -.500 | -.500 | -.500 |
| Mean of Estimate | | -.435 | -.435 | -.428 | -.425 |
| S.E. of Estimate | | .110 | .100 | .113 | .109 |
| MSE of Estimate | | .016 | .014 | .018 | .018 |
| Bias | | .065 | .065 | .072 | .075 |
| 95% CI Bias | | .06,.07 | .06,.07 | .06,.08 | .07,.08 |
| Percent of Bias | | 13% | 13% | 14% | 15% |
| Macro Parameter | $\gamma_{11}$ | | | | |
| The True Value | | .75 | .75 | .75 | .75 |
| Mean of Estimate | | .663 | .664 | .662 | .669 |
| S.E. of Estimate | | .136 | .129 | .147 | .146 |
| MSE of Estimate | | .026 | .024 | .029 | .028 |
| Bias | | -.087 | -.086 | -.088 | -.081 |
| 95% CI Bias | | -.10,-.07 | -.10,-.07 | -.10,-.08 | -.10,-.07 |
| Percent of Bias | | 12% | 12% | 12% | 11% |

Note
+ 400 replications were performed.
++ 800 replications were performed.

MSE and percent of bias for estimated $\gamma_{10}$ were .014 and 17%. The same estimates with RRSS were slightly smaller at .007 and 13% respectively. The results in Table 4-26 indicated that SPSS estimate of $\gamma_{10}$ and $\gamma_{11}$ had a slightly smaller bias and MSE for RRSS than for RRSL.

<u>The Effect of the Magnitude of the RRS Variance on the Estimated Standard Error of $\gamma$'s</u>

The mean, standard deviation, MSE and bias of estimated standard errors of the three macro parameters for both the estimation methods were compared in terms of small versus large variance of the random regression slope to the intercept variance (i.e., RRSS versus RRSL). See Tables 4-26, and 4-16 through 4-18.

The results indicated that RRSS and RRSL had no clear effects on the estimated standard error of the macro parameters.

In addition, the results again indicated clearly a smaller bias for $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$ with the VARCL estimates when compared with the SPSS estimates of the standard error of the same macro parameters in either condition (i.e., RRSS versus RRSL).

<u>Checking the Accuracy of the Variance-Covariance Component of the Random Effects Estimate Using VARCL</u>

One purpose of this study was to investigate the accuracy of the VARCL estimation method in estimating the variance-

Table 4-26.--The true standard error*, Mean, S.E., MSE, and bias for estimated standard error for macro parameters by the magnitude of random regression slope variance for the SPSS and VARCL estimation procedure.

| | Estimation Method | | | |
|---|---|---|---|---|
| | SPSS | VARCL | SPSS | VARCL |
| The Magnitude of Random Regression Slope Variance to the Intercept Variance | SMALL+ | SMALL+ | LARGE++ | LARGE++ |
| Macro Parameter $\gamma_{01}$ | | | | |
| The True standard error | .174 | .174 | .157 | .157 |
| Mean of Estimate | .101 | .166 | .101 | .157 |
| S.E. of Estimate | .044 | .032 | .043 | .033 |
| MSE of Estimate | .007 | .001 | .005 | .001 |
| Bias | -.073 | -.008 | -.056 | .001 |
| 95% CI Bias | -.08,-.07 | -.012,-.004 | -.06,-.05 | -.002,.002 |
| Percent of Bias | 42% | 5% | 36% | 0% |
| Macro Parameter $\gamma_{10}$ | | | | |
| The True standard error | .110 | .100 | .113 | .109 |
| Mean of Estimate | .090 | .096 | .091 | .106 |
| S.E. of Estimate | .041 | .043 | .041 | .038 |
| MSE of Estimate | .002 | .002 | .002 | .001 |
| Bias | -.020 | -.004 | -.022 | -.003 |
| 95% CI Bias | -.02,-.016 | -.01,.00 | -.024,-.02 | -.005,-.001 |
| Percent of Bias | 18% | 4% | 19% | 3% |
| Macro Parameter $\gamma_{11}$ | | | | |
| The True standard error | .136 | .129 | .147 | .146 |
| Mean of Estimate | .112 | .120 | .112 | .130 |
| S.E. of Estimate | .052 | .054 | .051 | .049 |
| MSE of Estimate | .003 | .003 | .004 | .003 |
| Bias | -.024 | -.009 | -.035 | -.016 |
| 95% CI Bias | -.03,-.02 | -.015,-.003 | -.04,-.03 | -.02,-.01 |
| Percent of Bias | 18% | 7% | 24% | 11% |

Note
+ 400 replications were performed.
++ 800 replications were performed.
*The true value was obtained from the observed standard error of the estimated macro parameters by the magnitude of random regression slope variance for the SPSS and VARCL estimation procedures shown in Table 4-25.

covariance components of the random effects (i.e., $\tau_{00}$, $\tau_{00}$, and $\tau_{01}$). In order to do this, several statistics were computed that took into account the difference between the VARCL estimated variance-covariance components of the random effects and its true value.

Another purpose of the study was to investigate the effect of different combined conditions on the estimates of the variance-covariance components of the random effects.

## Checking the Accuracy of the Estimated $\tau_{00}$ Obtained by the VARCL Estimation Method

Table 4-27 shows the true value, mean, standard error, MSE, and bias statistics for $\tau_{00}$ . These statistics were used to evaluate the accuracy of the VARCL estimation method of $\tau_{00}$ using the true values across six cells.

The results suggested that the VARCL estimates of the $\tau_{00}$ parameter were significantly biased. The negative sign of the bias implied that the VARCL underestimated $\tau_{00}$. In fact, the percent of the bias for the six cells ranged between 15% and 26%. When the bias was arranged from the smallest to the largest values, the magnitude of the bias was smaller as the number of units within each group increased. Similar trends were observed when the MSE was ordered in terms of magnitude.

## The Effect of n on $\tau_{00}$

The bias and MSE for estimated $\tau_{00}$ under different sample sizes, while holding the other factors fixed {i.e.,

Table 4-27.--The true value, Mean, S.E., MSE, and bias for estimated $\tau_{00}$ by cell identification for the VARCL estimation procedure*.

| Cell identification | (ND,n10,RRSS) | (ND,n10,RRSL) |
|---|---|---|
| The True Value | .995 | .85 |
| Mean of Estimate | .754 | .625 |
| S.E. of Estimate | .277 | .264 |
| MSE of Estimate | .134 | .120 |
| Bias | -.241 | -.225 |
| 95% CI Bias | -.28,-.20 | -.26,-.19 |
| Percent of Bias | 24% | 26% |
| **Cell identification** | **(ND,n60,RRSS)** | **(ND,n60,RRSL)** |
| The True Value | .995 | .85 |
| Mean of Estimate | .846 | .699 |
| S.E. of Estimate | .211 | .171 |
| MSE of Estimate | .067 | .052 |
| Bias | -.149 | -.151 |
| 95% CI Bias | -.18,-.12 | -.17,-.13 |
| Percent of Bias | 15% | 18% |
| **Cell identification** | **(TD,n10,RRSL)** | **(TD,n60,RRSL)** |
| The True Value | .85 | .85 |
| Mean of Estimate | .633 | .669 |
| S.E. of Estimate | .358 | .271 |
| MSE of Estimate | .174 | .106 |
| Bias | -.217 | -.181 |
| 95% CI Bias | -.27,-.17 | -.22,-.14 |
| Percent of Bias | 26% | 21% |

Note
* -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance.
RRSL-large magnitude of the random regression slope variance to the intercept variance.

(ND,n10,RRSS) versus (ND,n60,RRSS), (ND,n10,RRSL) versus (ND,n60,RRSL), and (TD,n10,RRSL) versus (TD,n60,RRSL)} were consistently smaller for n=60 than for n=10 (see Table 4-27). For example, comparing the (ND,n10,RRSL) versus (ND,n60,RRSL) cells, the bias and MSE for the $\tau_{00}$ parameter dropped from -.225 and .120 to -.151 and .052, respectively.

<u>The Effect of the Magnitude of the RRS Variance on the Estimated $\tau_{00}$</u>

The results indicated that the estimated $\tau_{00}$ parameter had a slightly smaller percent of bias for RRSS than for RRSL. This was done by comparing the (ND,n10,RRSS) versus (ND,n10,RRSL) and (ND,n60,RRSS) versus (ND,n60,RRSL) cells (see Table 4-27).

<u>The Effect of the Random Effects Distribution on the Estimated $\tau_{00}$</u>

The results showed that the estimated $\tau_{00}$ parameter had a smaller MSE for a normal distribution when compared to a t-distribution. By comparing the cells (ND,n10,RRSL) versus (TD,n10,RRSL) for n=10, the bias was about the same for both the random effects distributions (see Table 4-27). By comparing the cells (ND,n60,RRSL) versus (TD,n60,RRSL), with n=60, the bias was slightly smaller when the random effects distribution was normally distributed (-.151) as compared to when it was t-distributed (-.181) (see Table 4-27).

## Checking the Accuracy of the Estimated Standard Error of $\sqrt{\tau_{00}}$ obtained from VARCL Estimation procedure

Table 4-28 shows several statistics (i.e., true value, mean, standard error, MSE, and bias statistics) used to evaluate the accuracy of the VARCL estimation for the estimated standard error of $\sqrt{\tau_{00}}$ (notice that VARCL program reports the standard error for $\sqrt{\tau_{00}}$ rather than $\tau_{00}$) using the true values across six cells.

The results suggested that the VARCL estimates of the standard error for $\sqrt{\tau_{00}}$ were significantly biased, and the magnitude of the bias and MSE was smaller as the number of the units within each group increased.

## The Effect of n on Estimated Standard Error of $\sqrt{\tau_{00}}$

While holding the other factors fixed, the bias, the percent of bias, the standard deviation, and the MSE for estimated standard error for $\sqrt{\tau_{00}}$ under different sample sizes, were consistently smaller for n=60 than for n=10. see Table 4-28.

## The Effect of the Magnitude of the RRS Variance on Estimated Standard Error of $\sqrt{\tau_{00}}$

The results indicated a slightly larger percent of bias with RRSL than with RRSS for the estimated standard error of $\sqrt{\tau_{00}}$, after holding the other factors fixed. see Table 4-28.

Table 4-28.--The true value*, Mean, S.E., MSE, and bias for estimated standard error for $\sqrt{\tau_{00}}$ by cell identification for the VARCL estimation procedure**.

| Cell identification | (ND,n10,RRSS) | (ND,n10,RRSL) |
|---|---|---|
| The True Value | .085 | .084 |
| Mean of Estimate | .167 | .171 |
| S.E. of Estimate | .011 | .013 |
| MSE of Estimate | .007 | .008 |
| Bias | .082 | .087 |
| 95% CI Bias | .08,.084 | .085,.089 |
| Percent of Bias | 96% | 104% |
| **Cell identification** | **(ND,n60,RRSS)** | **(ND,n60,RRSL)** |
| The True Value | .096 | .083 |
| Mean of Estimate | .100 | .094 |
| S.E. of Estimate | .008 | .007 |
| MSE of Estimate | .000 | .000 |
| Bias | .004 | .011 |
| 95% CI Bias | .002,.006 | .009,.013 |
| Percent of Bias | 4% | 13% |
| **Cell identification** | **(TD,n10,RRSL)** | **(TD,n60,RRSL)** |
| The True Value | .121 | .117 |
| Mean of Estimate | .173 | .093 |
| S.E. of Estimate | .022 | .011 |
| MSE of Estimate | .003 | .001 |
| Bias | .052 | -.024 |
| 95% CI Bias | .048,.056 | -.026,-.022 |
| Percent of Bias | 43% | 21% |

Note
* -The true value was obtained form the standard deviation of the $\sqrt{\tau_{00}}$'s(i.e., square root of the true parameter, $\tau_{00}$) for each corresponding cell.
** -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance.
RRSL-large magnitude of the random regression slope variance to the intercept variance.

## The Effect of the Random Effects Distribution on Estimated Standard Error of $\sqrt{\tau_{00}}$

For n=10 the results in Table 4-28 indicated slightly a smaller bias and MSE but a slightly larger standard deviation on estimated standard error for $\sqrt{\tau_{00}}$ that had a t-distribution rather than one with a normal distribution. However, for n=60, the results indicated a smaller bias, MSE and standard deviation for the normal distribution than the t-distribution.

## Checking the Accuracy of Estimated $\tau_{11}$ obtained by the VARCL Estimation Method

Table 4-29 contains several statistics used to evaluate the accuracy of the VARCL estimation method of $\tau_{11}$ using the true values across six cells.

First, the results indicated that the VARCL estimates of $\tau_{11}$ were significantly biased, and the percent of bias was very large in (ND,n10,RRSS) cell. However, the magnitude of bias was reduced by increasing the sample size. In addition, the results also indicated that the size of MSE was clearly affected by the number of units within each group. The larger the sample size, the smaller the MSE.

## The Effect of n on the Estimated $\tau_{11}$

The results indicated that the MSE for estimated $\tau_{11}$ was smaller when n=60 than when n=10.

Table 4-29.--The true value, Mean, S.E., MSE, and bias for estimated $\tau_{11}$ by cell identification for the VARCL estimation procedure*.

| Cell identification | (ND,n10,RRSS) | (ND,n10,RRSL) |
|---|---|---|
| The True Value | .005 | .15 |
| Mean of Estimate | .057 | .134 |
| S.E. of Estimate | .086 | .168 |
| MSE of Estimate | .010 | .028 |
| Bias | .052 | -.016 |
| 95% CI Bias | .04,.06 | -.04,.01 |
| Percent of Bias | 1040% | 11% |
| **Cell identification** | **(ND,n60,RRSS)** | **(ND,n60,RRSL)** |
| The True Value | .005 | .15 |
| Mean of Estimate | .014 | .130 |
| S.E. of Estimate | .019 | .061 |
| MSE of Estimate | .000 | .004 |
| Bias | .009 | -.020 |
| 95% CI Bias | .007,.01 | -.03,-.01 |
| Percent of Bias | 180% | 13% |
| **Cell identification** | **(TD,n10,RRSL)** | **(TD,n60,RRSL)** |
| The True Value | .15 | .15 |
| Mean of Estimate | .131 | .120 |
| S.E. of Estimate | .170 | .066 |
| MSE of Estimate | .029 | .005 |
| Bias | -.019 | -.030 |
| 95% CI Bias | -.04,.005 | -.04,-.02 |
| Percent of Bias | 13% | 20% |

Note
* -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance.
RRSL-large magnitude of the random regression slope variance to the intercept variance.

The Effect of Magnitude of the RRS Variance on the Estimated $\tau_{11}$

The cells (ND,n10,RRSS) versus (ND,n10,RRSL) and (ND,n60,RRSS) versus (ND,n60,RRSL) were compared (see Table 4-29). The results indicated that the percent of bias of estimated $\tau_{11}$ was much larger with RRSL than with RRSS. See Table 4-29.

The Effect of Random Effects Distribution on Estimated $\tau_{11}$

By comparing the cells (ND,n10,RRSL) versus (TD,n10,RRSL) and (ND,n60,RRSL) versus (TD,n60,RRSL) the results for n=60 indicated that the $\tau_{11}$ parameter estimate had a slightly smaller bias, percent of bias, MSE, and standard deviation when the random effects had a normal distribution than when it had a t-distribution (see Table 4-29).

Checking the Accuracy of the Estimated Standard Error of $\sqrt{\tau_{11}}$ obtained from the VARCL Estimation Method

Table 4-30 shows the true value, mean, standard deviation, and bias of the VARCL estimated standard error for $\sqrt{\tau_{11}}$ across six cells, having 200 replications within each cell.

The Effect of n on Estimated Standard Error of $\sqrt{\tau_{11}}$

The results in Table 4-30 suggested that the VARCL estimates of the standard error for $\sqrt{\tau_{11}}$ were significantly biased, and the percent of bias was very large specially for

Table 4-30.--The true value*, Mean, S.E., MSE, and bias for estimated standard error for $\sqrt{\tau_{11}}$ by cell identification for the VARCL estimation procedure**.

| Cell identification | (ND,n10,RRSS) | (ND,n10,RRSL) |
|---|---|---|
| The True Value | .006 | .032 |
| Mean of Estimate | .311 | .309 |
| S.E. of Estimate | .165 | .141 |
| MSE of Estimate | .120 | .096 |
| Bias | .305 | .277 |
| 95% CI Bias | .28,.33 | .26,.30 |
| Percent of Bias | 5083% | 865% |
| **Cell identification** | **(ND,n60,RRSS)** | **(ND,n60,RRSL)** |
| The True Value | .006 | .034 |
| Mean of Estimate | .134 | .076 |
| S.E. of Estimate | .095 | .013 |
| MSE of Estimate | .025 | .002 |
| Bias | .128 | .042 |
| 95% CI Bias | .11,.14 | .04,.044 |
| Percent of Bias | 2133% | 124% |
| **Cell identification** | **(TD,n10,RRSL)** | **(TD,n60,RRSL)** |
| The True Value | .051 | .054 |
| Mean of Estimate | .300 | .079 |
| S.E. of Estimate | .169 | .015 |
| MSE of Estimate | .091 | .001 |
| Bias | .249 | .025 |
| 95% CI Bias | .23,.27 | .023,.027 |
| Percent of Bias | 488% | 46% |

Note
* -The true value was obtained form the standard deviation of the $\sqrt{\tau_{11}}$'s(i.e., square root of the true parameter, $\tau_{11}$) for each corresponding cell.
** -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance.
RRSL-large magnitude of the random regression slope variance to the intercept variance.

small sample sizes. In addition, the magnitude of the bias, standard deviation, and MSE were reduced as the sample size within each group increased, holding the other factors as fixed.

### The Effect of the Magnitude of the RRS Variance on the Estimated Standard Error of $\sqrt{\tau_{11}}$

Comparing (ND,n10,RRSS) versus (ND,n10,RRSL) and (ND,n60,RRSS) versus (ND,n60,RRSL), the results suggested that increasing the magnitude of RRS variance to the intercept variance led to a smaller bias, MSE, and standard deviation estimate of the standard error for $\sqrt{\tau_{11}}$.

### The Effect of the Random Effects Distribution on the Estimated Standard Error of $\sqrt{\tau_{11}}$

The results indicated a smaller percent of bias for the estimated standard error for $\sqrt{\tau_{11}}$ when the random effects had a t-distributed than when it was normally distributed, holding other factors as fixed.

### Checking the Accuracy of the Estimated $\tau_{01}$ Obtained by the VARCL Estimation Method

Table 4-31 contains several statistics used to evaluate the accuracy of the VARCL estimation method of $\tau_{01}$ using the true value, across six cells.

Table 4-31.--The true value, Mean, S.E., MSE, and bias for estimated $\tau_{01}$ by cell identification for the VARCL estimation procedure*.

| Cell identification | (ND,n10,RRSS) | (ND,n10,RRSL) |
|---|---|---|
| The True Value | .000 | .000 |
| Mean of Estimate | .033 | .030 |
| S.E. of Estimate | .151 | .158 |
| MSE of Estimate | .024 | .026 |
| Bias | .033 | .030 |
| 95% CI Bias | .01,.05 | .01,.05 |
| Percent of Bias | NA | NA |
| **Cell identification** | **(ND,n60,RRSS)** | **(ND,n60,RRSL)** |
| The True Value | .000 | .000 |
| Mean of Estimate | .019 | .002 |
| S.E. of Estimate | .063 | .078 |
| MSE of Estimate | .004 | .006 |
| Bias | .019 | .002 |
| 95% CI Bias | .011,.03 | -.009,.014 |
| Percent of Bias | NA | NA |
| **Cell identification** | **(TD,n10,RRSL)** | **(TD,n60,RRSL)** |
| The True Value | .000 | .000 |
| Mean of Estimate | .020 | .013 |
| S.E. of Estimate | .184 | .084 |
| MSE of Estimate | .034 | .007 |
| Bias | .020 | .013 |
| 95% CI Bias | -.005,.05 | .001,.02 |
| Percent of Bias | NA | NA |

Note
* -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance.
RRSL-large magnitude of the random regression slope variance to the intercept variance.
NA - Not Applicable.

## The Effect of n on Estimated $\tau_{01}$

The results in Table 4-31 indicated that the size of bias, standard deviation, and MSE were clearly affected by the number of units within each group. The larger the sample size, the smaller the bias, standard deviation and MSE for the estimated $\tau_{01}$.

## The Effect of the Magnitude of the RRS Variance on the Estimated $\tau_{01}$

By comparing the cell (ND,n10,RRSS) versus (ND,n10,RRSL), it was found that the effect of magnitude of the random regression slope was very small when n=10. See Table 4-31. Similarly, by comparing the cell (ND,n60,RRSS) versus (ND,n60,RRSL) the bias for estimated $\tau_{01}$ was smaller for RRSL (.002) than for RRSS (.019) when n=60.

## The Effect of Random Effects Distribution on the Estimated $\tau_{01}$

By comparing the cell (ND,n60,RRSL) versus (TD,n60,RRSL) (see Table 4-31), it was observed that the $\tau_{01}$ parameter estimate had a smaller bias, standard deviation, and MSE when the random effects had a normal distribution as compared to a t-distribution, for n=60. The type of the random effects distribution had no clear effect when n=10.

Checking the Accuracy of the Estimated Standard Error of $\tau_{01}$ Obtained by the VARCL Estimation

Table 4-32 shows the true value, mean, standard deviation, and bias of VARCL estimated standard error for $\tau_{01}$ across six cells, having 200 replications within each cell.

The Effect of n on Estimated Standard Error of $\tau_{01}$

The results in Table 4-32 showed that the VARCL estimates of the standard error for $\tau_{01}$ were significantly biased, with a large percent of bias. However, the magnitude of bias, standard deviation, and MSE became smaller as the sample size within each group increased from 10 to 60, holding the other factors as fixed.

The Effect of the Magnitude of the RRS Variance on the Estimated Standard Error of $\tau_{01}$

The results indicated that for n=10 and n=60, increasing the magnitude of RRS variance to the intercept variance led to a smaller bias and MSE estimate of the standard error for $\tau_{01}$.

The Effect of the Random Effects Distribution on the Estimated Standard Error of $\tau_{01}$

The results in Table 4-32 indicated that for both n=10 and n=60, having a normal distribution of the random effects led to a slightly smaller bias, percent of bias, standard deviation, and MSE estimated standard error for $\tau_{01}$ than having a t-distribution of the random effects.

Table 4-32.--The true value*, Mean, S.E., MSE, and bias for estimated standard error for $\tau_{01}$ by cell identification for the VARCL estimation procedure**.

| Cell identification | (ND,n10,RRSS) | (ND,n10,RRSL) |
|---|---|---|
| The True Value | .008 | .045 |
| Mean of Estimate | .162 | .162 |
| S.E. of Estimate | .023 | .024 |
| MSE of Estimate | .024 | .014 |
| Bias | .154 | .117 |
| 95% CI Bias | .15,.16 | .11,.12 |
| Percent of Bias | 1925% | 260% |
| **Cell identification** | **(ND,n60,RRSS)** | **(ND,n60,RRSL)** |
| The True Value | .010 | .045 |
| Mean of Estimate | .052 | .064 |
| S.E. of Estimate | .007 | .011 |
| MSE of Estimate | .002 | .000 |
| Bias | .042 | .019 |
| 95% CI Bias | .042,.042 | .017,.02 |
| Percent of Bias | 420% | 42% |
| **Cell identification** | **(TD,n10,RRSL)** | **(TD,n60,RRSL)** |
| The True Value | .041 | .039 |
| Mean of Estimate | .162 | .062 |
| S.E. of Estimate | .026 | .012 |
| MSE of Estimate | .015 | .001 |
| Bias | .121 | .023 |
| 95% CI Bias | .12,.13 | .021,.025 |
| Percent of Bias | 295% | 59% |

Note
* -The true value was obtained form the standard deviation of the $\tau_{01}$'s (i.e., square root of the true parameter, $\tau_{01}$) for each corresponding cell.
** -200 replications were performed within each cell.
ND -normal distribution of the random effects.
TD -t-distribution of the random effects.
n10 -10 subjects within each group.
n60 -60 subjects within each group.
RRSS-small magnitude of the random regression slope variance to the intercept variance.
RRSL-large magnitude of the random regression slope variance to the intercept variance.

## Summary

A summary statistics of the key results that were discussed in this chapter is presented in Table 4-33.

Table 4-33.--A summary of several statistics for different parameter by the SPSS and the VARCL estimation procedure.

| Parameter | Bias | | Consistency | | MSE | | Type I Error Rate | |
|---|---|---|---|---|---|---|---|---|
| | Estimation Method | | | | | | | |
| | SPSS | VARCL | SPSS | VARCL | SPSS | VARCL | SPSS | VARCL |
| $\gamma_{oo}$ | Yes | Yes | No | No | ND | ND | NA | NA |
| $\gamma_{o1}$ | Yes | Yes | No | No | ND | ND | H | - |
| $\gamma_{1o}$ | Yes | Yes | No | No | ND | ND | H | - |
| $\gamma_{11}$ | Yes | Yes | No | No | ND | ND | H | - |
| S.E.($\gamma_{o1}$) | Yes | No | No | Yes | - | S | - | - |
| S.E.($\gamma_{1o}$) | Yes | No | No | Yes | ND | ND | - | - |
| S.E.($\gamma_{11}$) | Yes | No | No | Yes | ND | ND | - | - |
| $\tau_{oo}$ | NA | Yes | NA | Yes | NA | .109 | NA | - |
| $\tau_{11}$ | NA | Yes | NA | No | NA | .013 | NA | - |
| $\tau_{o1}$ | NA | Yes | NA | Yes | NA | .017 | NA | - |
| S.E.($\sqrt{\tau_{oo}}$) | NA | Yes | NA | Yes | NA | .003 | NA | - |
| S.E.($\sqrt{\tau_{11}}$) | NA | Yes | NA | Yes | NA | .056 | NA | - |
| S.E.($\tau_{o1}$) | NA | Yes | NA | Yes | NA | .009 | NA | - |

Note
MSE -An average Mean Square Error across six cells.
ND -No difference between the two estimation methods.
NA -Not applicable.
H -Higher than the other estimation method.
S -Smaller than the other estimation method.

# CHAPTER V

# CONCLUSION

## Introduction

This chapter presents the conclusions of the analyzes of the study. The chapter begins by first presenting the conclusions based on the real data analysis. This will be followed by the conclusions based on the simulated data analysis. The implications of the findings will then be addressed. This is followed by a discussion of the consequences of the real and simulated data analysis conclusions. The final section of this chapter will present some suggestions for future research.

## Conclusions Based on the Real Data Analysis

The conclusions based on running the SPSS, VARCL, and MULTILOGIT estimation methods on real data are as follows:

(1) The regression coefficient estimates for the within-school variable for the SPSS and VARCL approaches were close, while that the MULTILOGIT estimate had a larger absolute value than both the SPSS and VARCL approaches. However, there appears to be no consistent pattern with regard to the regression coefficient estimates for the school-level variables.

(2) The estimated standard error of the regression coefficient for both the within- and between-school variables for the VARCL and MULTILOGIT approaches using the real data were close. However, the MULTILOGIT estimates were slightly larger than the VARCL estimates.

(3) The results of real data analysis also indicated that the magnitude of the VARCL and MULTILOGIT estimates of the standard error of the regression coefficient were much larger than the SPSS estimates.

(4) The variance-covariance components of the random effects estimate of MULTILOGIT and VARCL using the real data were close. However, the MULTILOGIT estimates were generally larger in absolute values than the VARCL estimates.

## Conclusions Based on the Simulated Data Analysis

The following conclusions were based on running the SPSS and VARCL programs estimation procedures on simulated data that were generated for the multilevel logistic regression model (a random effects model with binary outcomes):

(1) Both the VARCL and SPSS estimates of $\gamma$'s were found to be significantly biased. The percentages of biased ranged between 10% and 17% lower than the true values. The VARCL and SPSS estimates of $\gamma_{00}$, $\gamma_{01}$ , $\gamma_{10}$ and $\gamma_{11}$ parameter were found to be approximately equal for different statistics (i.e., mean, standard deviation, MSE, and bias). For both estimation methods, increasing the number of units within each group (n) resulted in slightly increasing the bias of the estimated

macro parameters. However, increasing n led to a slightly smaller MSE of the $\gamma_{oo}$, $\gamma_{o1}$, $\gamma_{1o}$, and $\gamma_{11}$ estimates for both the VARCL and SPSS estimation methods. This reduction in MSE is caused by the smaller magnitude of the standard deviation of the estimated macro parameters as a result of increasing the sample size. There was also no clear effect of the random effects distributions (i.e., ND versus TD) on all four macro parameters for both estimation procedures. Finally, the VARCL estimate of the macro, $\gamma_{1o}$ and $\gamma_{11}$, parameter estimates had a slightly smaller bias and MSE for RRSS (having a small magnitude of random regression slope variance in contrast to the intercept variance) as compared to RRSL (having a large magnitude of random regression slope variance in contrast to the intercept variance). The magnitude of the random regression slope variance appeared to have no clear effect on the VARCL estimate of $\gamma_{oo}$ and $\gamma_{o1}$ parameters. In addition, the results also indicated that the SPSS estimate of $\gamma_{1o}$ and $\gamma_{11}$, had a slightly smaller bias and MSE for RRSS as compared to RRSL. While there was no clear effect on the SPSS estimation of the macro parameters, $\gamma_{oo}$ and $\gamma_{o1}$.

Therefore, under the random effects model for binary outcomes, the VARCL estimates of the macro parameter was significantly biased and inconsistent. A similar result was obtained for the SPSS estimates of the macro parameters. In fact, the results in Table 4-13 indicated that the SPSS of $\gamma_{oo}$ (from simple random effects model for binary outcome) moved further away from the true value (i.e., the magnitude of bias increases) as the true value for $\tau_{00}$ increases. This is because the generated data under the random effects model is different from the SPSS model assumptions.

(2) On the average, the estimated standard errors of $\gamma$'s for VARCL were larger than the SPSS estimate of the standard errors. And the SPSS estimates of the standard error of macro parameters were clearly significantly biased, while the VARCL estimates of the standard error of macro parameter were unbiased. This is due to the larger estimates of the standard errors for $\gamma$'s of VARCL when compared to SPSS. In addition, in for both estimation methods, the random effects distributions and the magnitude of the random regression slope variance had no clear effect on the estimated standard errors of the estimated macro parameters. Increasing the sample size resulted in slightly smaller standard deviation and MSE of the VARCL estimates of the standard error of the three macro parameters (i.e., $\gamma_{01}$, $\gamma_{10}$ and $\gamma_{11}$). However, with the SPSS program, increasing the sample size resulted in estimates of the standard error of the three macro parameters that were slightly larger in bias and MSE.

(3) The probability of type I error rate under a true null hypothesis tests of the macro parameters $\gamma$'s were much smaller for the VARCL than the SPSS program. However, both estimation methods gave unacceptable high type I error rates (i.e., $p > .05$).

(4) The VARCL estimates of $\tau_{00}$, $\tau_{11}$ and $\tau_{01}$ parameters were significantly biased and underestimated the true values.

However, the magnitude of the bias and MSE was reduced as the number of units within each group increased. The magnitude of the regression slope variance (i.e., RRSL vs. RRSS) had no clear effect on $\tau_{00}$ and $\tau_{01}$. Except for $\tau_{11}$, the percentage of bias were smaller for RRSL when compared to RRSS. Finally, the results also indicated that the estimated variance-covariance components of the random effects parameter had a slightly smaller bias, MSE, and standard deviation when the random effects had a normal distribution than when it had a t-distribution, explicitly for large n.

(5) The VARCL estimates of the standard error for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{11}}$, and $\tau_{01}$ were significantly biased. However, the magnitude of the bias, standard deviation, and MSE were reduced as the sample size within each group increased from 10 to 60 (i.e., consistent). Increasing the magnitude of random regression slope variance to the intercept variance led to a smaller percentage of bias, MSE, and standard deviation estimate of the standard error for $\sqrt{\tau_{11}}$, and $\tau_{01}$. However, there was slightly smaller percentage of bias for RRSS when compared to RRSL for the estimated standard error of $\sqrt{\tau_{00}}$. Finally, a large n (i.e., n=60) for the normally distributed random effects resulted in a slightly smaller bias, standard deviation, and MSE of the estimated standard error for $\sqrt{\tau_{00}}$, and $\tau_{01}$ when this was compared to the t-distributed random effects.

## Implications of the Findings

The first part of this chapter addressed the statistical accuracy of the computer estimation programs on real and simulated data. However, this section will address the implications of the findings by identifying the limitations, the advantages, and the disadvantages of running these programs. The usefulness of some of the reported statistics for the investigators in making critical educational decisions will also be discussed.

The SPSS program estimation for a random effects model for binary outcomes indicated several disadvantages:

(1) The SPSS estimates of $\gamma$'s were found to be significantly biased and inconsistent. The estimates underestimated the true value.

(2) the SPSS estimates of the standard error of macro parameters were significantly biased and inconsistent. Increasing the sample size resulted in SPSS estimates of the standard error of the three macro parameters having a larger bias and MSE.

(3) The SPSS estimates gave a large probability of type I error rate under a true null testing the macro parameters, $\gamma$'s.

Similarly, there were some disadvantages in the using the current VARCL program:

(1) The VARCL estimates of $\gamma$'s were found to be significantly biased and inconsistent. The estimates underestimated the true value.

(2) The VARCL estimates of the standard error of macro parameter proved to be unbiased and consistent. This meant that increasing the sample size resulted in the VARCL estimates of the standard error of the three macro parameters having a smaller bias and MSE.

(3) The VARCL estimates gave a small probability of type I error rate under a true null testing of the macro parameters, $\gamma$'s, relative to the SPSS estimates. However, the VARCL type I error rate was not small enough to be acceptable (i.e., $p > .05$).

(4) The VARCL estimates of the $\tau_{00}$, $\tau_{11}$ and $\tau_{01}$ parameters were significantly biased, and underestimated the true values. However, the magnitude of the bias and MSE were reduced as the number of units within each group increased (i.e., they were consistent).

(5) The VARCL estimates of the standard error for $\sqrt{\tau_{00}}$, $\sqrt{\tau_{11}}$, and $\tau_{01}$ were significantly biased. However, the magnitude of bias, standard deviation, and MSE were reduced as the sample size within each group increased. In other words they were consistent.

The simulation study demonstrated that using the standard logistic regression estimation procedure for multilevel data with binary outcomes could lead to misleading results and conclusions. This is because the standard logistic regression estimates were found to be significantly biased and inconsistent for both the $\gamma$'s and the standard error of $\gamma$'s.

The following explanation is given for the bias that exists when the fixed model is used to develop an estimate for the intercept (i.e., the random intercept logistic regression model, see model A on page 86, and the SPSS and VARCL estimates in Table 4-13) in a random effects model. Consider the logistic function

$$y = f(\alpha) = \frac{e^{\alpha}}{1+ e^{\alpha}} = \frac{1}{1+ e^{-\alpha}}$$

The first derivative of the logistic function (refer to Appendix 5-1) is as follows

$$f'(\alpha) = \frac{e^{-\alpha}}{(1+ e^{-\alpha})^2}$$

The second derivative of the logistic function is given as

$$f''(\alpha) = \frac{e^{-\alpha}(e^{-2\alpha} -1)}{(1+ e^{-\alpha})^4} .$$

Let the a random variable, $\alpha$, be expressed as follows

$$\alpha = \gamma + u$$

where

$u \sim N(0,\sigma^2)$ and

$\gamma$ is a constant (i.e., intercept).

Expanding $f(\alpha)$ (i.e., Taylor expansion) about $\gamma$ up to quadratic terms,

$$y = f(\alpha) \doteq f(\gamma) + f'(\gamma)(\alpha - \gamma) + \frac{f''(\gamma)}{2}(\alpha - \gamma)^2$$

Hence

$$y = f(\alpha) \doteq f(\gamma) + f'(\gamma)(u) + \frac{f''(\gamma)}{2}(u)^2$$

note that $E(u) = 0$, $E(u^2) = \sigma^2$, and

$$E(y) \doteq f(\gamma) + \frac{f''(\gamma)}{2}\sigma^2$$

For example, let $\gamma = -1.80$, $\sigma^2 = .85$. Thus E(y) is given as E(y) = .1419 + (.0872 / 2) (.85) = .1790. The logit (.1790) = -1.52, where as logit (.1419) = -1.80.

Therefore, if E(y) is estimated by an unbiased estimate, the logit of this estimate will be about -1.52 whereas the intercept, $\gamma$, is -1.80. In fact, the similarity of the VARCL and the SPSS estimates of $\gamma$'s (see Table 4-13) makes it highly likely that a similar reasoning will explain the bias of the VARCL estimates of $\gamma$'s.

The explanation for the significantly biased estimates of the standard error of the regression parameters for the standard logistic regression estimation method, in case of multilevel data, is attributed to ignoring the parameter variance of the single level model in its estimate of the standard error of the regression parameters. The multilevel approaches account for both the parameter variance and sampling variance in its estimate for the standard error of the regression parameters.

Therefore, caution should be exercised when studying multilevel data with binary outcomes using the standard single logistic regression estimation procedure (i.e., the SPSS program) instead of the multilevel logistic regression estimation procedure. This is because of the high probability of a type I error for the standard single logistic regression estimation method (see Tables 4-19 and 4-20). This error was due to the liberal t-statistic values, caused partly by the small standard error estimates for the regression coefficients, and partly by the significantly biased estimates of $\gamma$'s when assuming the single-level logistic regression model by using the standard logistic regression estimation method. In addition, the VARCL type I error rate (under a true null hypothesis, $H_0$: $\gamma = \gamma_t$) was not small enough to be ignored (i.e., p >.05). This was because the VARCL estimates of $\gamma$'s were found to be significantly biased, inconsistent, and underestimated the true values.

The results of the real data showed that: (a) the estimated regression coefficient for the MULTILOGIT had a larger absolute value than the VARCL estimate, (b) the estimated standard errors of the regression coefficient for MULTILOGIT were slightly larger than the VARCL estimates, and (c) the MULTILOGIT estimates for variance-covariance components of the multilevel logistic regression model were generally larger absolute value than VARCL estimates. Based on the knowledge that the current VARCL program underestimated both the (a) the macro parameters, and (b) variance-covariance components of the random effects (Longford, 1992), the MULTILOGIT program may be more efficient program than the VARCL.

However, there were several reasons that made operating the VARCL program more attractive than the MULTILOGIT program. These are summarized as follows:

(1) The MULTILOGIT program had a limit in the number of micro and macro variables that could be included in an analysis. No such limitation was indicated by the VARCL program.

(2) The MULTILOGIT program also had a limit of 59 groups (or schools) that could be used in the analysis. Again no such limitation exists for the VARCL program.

(3) The MULTILOGIT program proved to be inconvenient to operate. This was essentially because the coefficient input file required the researcher to provide the estimates of the classical within-group logistic regression coefficients for each school in the analysis. On the other hand, the VARCL program generated its own initial estimates.

(4) The MULTILOGIT program model specification always assumed that all the micro regression (intercept and slope) were random coefficients. The VARCL program, however, had the option to assume fixed or random regression coefficients among schools. In fact, the ability of the VARCL program to test the variance-covariance components of the random effects is

critical for the investigator in deciding whether to assume fixed or random regression coefficients. This facility is not available for the MULTILOGIT program user.

(5) The inability of the educational researcher to run the more complicated MLRM with the MULTILOGIT program (i.e., by including more covariates in the within-school model) was due to the small number of students within each school. This insensitivity to the small number of subjects within each group was not observed with the VARCL program.

(6) It was found to be financially very expensive to run the MULTILOGIT program at the University of Michigan Mainframe Computer Center. The personal computer version of the MULTILOGIT program is presently unavailable.

## The Consequences of the Conclusions of the Real and Simulated Data Analyses

The argument in the last section indicated several statistical disadvantages in using the SPSS program for the random effects model having binary outcomes. Similarly, the VARCL indicated some disadvantages in estimation the (a) the macro parameters, and (b) variance-covariance components of the random effects. Therefore, based on the parameter estimation, the MULTILOGIT program may be more efficient than the VARCL programs. However, There are several reasons that has been indicated in the last section made operating the MULTILOGIT program very restrictive. Thus, if one were able to account for the existing bias in the current the VARCL program, it would be perhaps more advantageous to choose the VARCL program over MULTILOGIT program. An exception would be if the researcher can accept the disadvantages and limitations of the MULTILOGIT program (i.e., cost, sample size within each group, limitation in the number of covariates, limitation in the number of groups, inability to provide standard logistic regression estimates coefficient of each group, inability of the researcher to statistically decide whether to fixed or assumed the random regression coefficient among groups).

## Suggestions for Future Research

This study suggests that future research in developing an new program that accounts for the disadvantages in both the VARCL and MULTILOGIT programs. In addition, to overcoming the above disadvantages, the new program should efficient for small number of subjects within each group. This would represent a more realistic educational research situation. In fact, concerns regard a small number of subjects within each group on parameter estimations for binary outcome were indicated by Longford (1992). The MULTILOGIT program required a large number of subjects within each group in order to run.

Finally, the new program may also consider a more simplified model due to the nature of the outcome variable (i.e., binary outcome). Like one having a random intercept and a fixed regression coefficient slope model (**model I**) rather than one having a random intercept and a random regression coefficient slope model (**model II**). In fact, several researchers (Chamberlain, 1980; Korn and Whittemore, 1979) have recommended this for normally distributed outcomes. This would make the estimation procedure of model I less complex than the model II. Raudenbush (1988) has also pointed out the advantage for this by stating:

> A random intercept model has two computational advantages: (a) the number of microcoefficients reduces to one per group; and, therefore, (b) the variance-covariance matrix of the random effects (T in our notation) becomes diagonal, which simplifies estimation formulas (p. 106).

Similarly, the simplicity of the random intercept model was also indicated by Wong and Mason (1985). In addition to these advantages, Shigemasu (1976) indicated concerns regarding the cost of computation for using model II saying that "... the model (i.e., model I) is expected to reduce substantially the cost computation" (p. 158).

# APPENDICES

# APPENDIX 3-1

## AN EXAMPLE OF VARCL PROGRAM "BASIC INFORMATION FILE" SPECIFIED FOR THIS STUDY

---

```
SIMULATED LRM FOR 60 SCHOOLS HAVING 10 STUDENTS IN EACH SCHOOL
    2
   600      60
     3       1
    90      10       4
student
school
d:\vac_02dt.001
(f8.5,1x,f8.5,1x,f8.5)
d:\schvar.dat
(9x,F8.5)
dep          2
xij          1
INTXiZ2      1
Z2j          1
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
   10    10    10    10    10    10
```

## APPENDIX 3-2

### AN EXAMPLE OF MULTILOGIT PROGRAM "COMMAND FILE" SPECIFIED FOR THIS STUDY

---

```
\ This is the first heading line.
\ This is the second heading line.
* This is the first line of comments.
* This is the second line of comments.
* This is the last line of comments.
   2    0.01   0.001  10    0.01   0.001   0
   9   1sx0t:stu59mts.DAT
   6   1sx0t:sch59mts.DAT
   2    sx0t:model3.DAT
-TEMP
  59   9   2   1   6
school01    school02    school03    school04    school05    school06
school07    school08    school09    school10    school11    school12
school13    school14    school15    school16    school17    school18
school19    school20    school21    school22    school23    school24
school25    school26    school27    school28    school29    school30
school31    school32    school33    school34    school35    school36
school37    school38    school39    school40    school41    school42
school43    school44    school45    school46    school47    school48
school49    school50    school51    school52    school53    school54
school55    school56    school57    school58    school59
repeat       subject
intercept     SES
   0   0

gammaoo

gammao1
```

# APPENDIX 3-3

## THE DESCRIPTIVE STATISTICS FOR THE REAL DATA AT BOTH THE STUDENT- AND SCHOOL-LEVEL

Descriptive statistics For 59 schools used in real data analysis.

School level

| Variable | Mean | Std Dev | Minimum | Maximum | N |
|---|---|---|---|---|---|
| URB_RRL | .00 | .44 | -.25 | .75 | 59 |
| CENTRAL | .00 | .43 | -.24 | .76 | 59 |
| NORTH | .00 | .38 | -.17 | .83 | 59 |
| SOUTH | .00 | .39 | -.19 | .81 | 59 |
| BANGKOK | .07 | .25 | .00 | 1.00 | 59 |
| MSES | .00 | .46 | -.90 | 1.81 | 59 |

URB_RRL URBAN/RURAL AREA

| Value Label | Value | Frequency | Percent | Valid Percent | |
|---|---|---|---|---|---|
| | -.25 | 44 | 74.6 | 74.6 | 74.6 |
| | .75 | 15 | 25.4 | 25.4 | 100.0 |
| | Total | 59 | 100.0 | 100.0 | |

CENTRAL

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.24 | 45 | 76.3 | 76.3 | 76.3 |
| | .76 | 14 | 23.7 | 23.7 | 100.0 |
| | Total | 59 | 100.0 | 100.0 | |

NORTH

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.17 | 49 | 83.1 | 83.1 | 83.1 |
| | .83 | 10 | 16.9 | 16.9 | 100.0 |
| | Total | 59 | 100.0 | 100.0 | |

SOUTH

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.19 | 48 | 81.4 | 81.4 | 81.4 |
| | .81 | 11 | 18.6 | 18.6 | 100.0 |
| | Total | 59 | 100.0 | 100.0 | |

Mean SES (MSES)

| | | | | | |
|---|---|---|---|---|---|
| Mean | -.001 | Std err | .060 | Median | -.146 |
| Mode | -.902 | Std dev | .460 | Variance | .212 |
| Kurtosis | 3.982 | S E Kurt | .613 | Skewness | 1.672 |
| S E Skew | .311 | Range | 2.711 | Minimum | -.902 |
| Maximum | 1.808 | Sum | -.060 | | |
| Valid cases | 59 | Missing cases | 0 | | |

Student level: 1244 student were involved in this analysis.

| Variable | Mean | Std Dev | Minimum | Maximum | Valid N |
|---|---|---|---|---|---|
| SCHOOLID | 98078.12 | 54720.88 | 10101.00 | 180550.0 | 1244 |
| URB_RRL | .06 | .46 | -.25 | .75 | 1244 |
| CENTRAL | .00 | .43 | -.24 | .76 | 1244 |
| NORTH | .03 | .40 | -.17 | .83 | 1244 |
| SOUTH | -.03 | .37 | -.19 | .81 | 1244 |
| MSES | .06 | .48 | -.90 | 1.81 | 1244 |
| SEX | .00 | .50 | -.50 | .50 | 1244 |
| DIALECT | .00 | .50 | -.49 | .51 | 1244 |
| LUNCH | .00 | .37 | -.84 | .16 | 1244 |
| SES | .00 | .68 | -1.72 | 3.28 | 1244 |
| SCPPED1 | .00 | 1.00 | -1.09 | .91 | 1244 |
| BRAKFAST | .00 | .39 | -.81 | .19 | 1244 |
| REP1 | .15 | .36 | .00 | 1.00 | 1244 |
| SEX_MSES | .00 | .38 | -.94 | 1.77 | 1244 |
| SCP_MSES | .00 | .43 | -.86 | 1.59 | 1244 |
| URB_SES | .00 | .35 | -.89 | 2.32 | 1244 |

URB_RRL URBAN/RURAL AREA

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.25 | 854 | 68.6 | 68.6 | 68.6 |
| | .75 | 390 | 31.4 | 31.4 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

CENTRAL

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.24 | 946 | 76.0 | 76.0 | 76.0 |
| | .76 | 298 | 24.0 | 24.0 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

NORTH

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.17 | 999 | 80.3 | 80.3 | 80.3 |
| | .83 | 245 | 19.7 | 19.7 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

SOUTH

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.19 | 1046 | 84.1 | 84.1 | 84.1 |
| | .81 | 198 | 15.9 | 15.9 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

Mean SES (MSES)

| | | | | | |
|---|---|---|---|---|---|
| Mean | .056 | Std err | .014 | Median | -.116 |
| Mode | .537 | Std dev | .481 | Variance | .231 |
| Kurtosis | 2.995 | S E Kurt | .139 | Skewness | 1.508 |
| S E Skew | .069 | Range | 2.711 | Minimum | -.902 |
| Maximum | 1.808 | Sum | 70.096 | | |
| Valid cases | 1244 | Missing cases | 0 | | |

SEX

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.50 | 621 | 49.9 | 49.9 | 49.9 |
| | .50 | 623 | 50.1 | 50.1 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

DIALECT

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.49 | 636 | 51.1 | 51.1 | 51.1 |
| | .51 | 608 | 48.9 | 48.9 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

LUNCH DO STUDENT HAVE LUNCH DAILY

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.84 | 200 | 16.1 | 16.1 | 16.1 |
| | .16 | 1044 | 83.9 | 83.9 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

SES

| | | | | | |
|---|---|---|---|---|---|
| Mean | -.004 | Std err | .019 | Median | -.213 |
| Mode | -.229 | Std dev | .680 | Variance | .463 |
| Kurtosis | 5.528 | S E Kurt | .139 | Skewness | 1.993 |
| S E Skew | .069 | Range | 5.003 | Minimum | -1.719 |
| Maximum | 3.283 | Sum | -4.410 | | |
| Valid cases | 1244 | Missing cases | 0 | | |

SCPPED1

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -1.09 | 569 | 45.7 | 45.7 | 45.7 |
| | .91 | 675 | 54.3 | 54.3 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

BRAKFAST

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| | -.81 | 234 | 18.8 | 18.8 | 18.8 |
| | .19 | 1010 | 81.2 | 81.2 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

REP1 EVER REPETITION

| Value Label | Value | Frequency | Percent | Valid Percent | Cum Percent |
|---|---|---|---|---|---|
| NEVER | .00 | 1052 | 84.6 | 84.6 | 84.6 |
| EVER | 1.00 | 192 | 15.4 | 15.4 | 100.0 |
| | Total | 1244 | 100.0 | 100.0 | |

SEX_MSES

| | | | | | |
|---|---|---|---|---|---|
| Mean | -.002 | Std err | .011 | Median | -.040 |
| Mode | -.040 | Std dev | .378 | Variance | .143 |
| Kurtosis | 10.363 | S E Kurt | .139 | Skewness | 2.690 |
| S E Skew | .069 | Range | 2.711 | Minimum | -.942 |
| Maximum | 1.768 | Sum | -3.094 | | |
| Valid cases | 1244 | Missing cases | 0 | | |

SCP_MSES

| | | | | | |
|---|---|---|---|---|---|
| Mean | -.001 | Std err | .012 | Median | -.057 |
| Mode | .317 | Std dev | .432 | Variance | .186 |
| Kurtosis | 3.069 | S E Kurt | .139 | Skewness | 1.358 |
| S E Skew | .069 | Range | 2.452 | Minimum | -.864 |
| Maximum | 1.588 | Sum | -.788 | | |
| Valid cases | 1244 | Missing cases | 0 | | |

URB_SES

| | | | | | |
|---|---|---|---|---|---|
| Mean | .000 | Std err | .010 | Median | -.043 |
| Mode | -.043 | Std dev | .351 | Variance | .123 |
| Kurtosis | 12.876 | S E Kurt | .139 | Skewness | 2.643 |
| S E Skew | .069 | Range | 3.209 | Minimum | -.887 |
| Maximum | 2.322 | Sum | .168 | | |
| Valid cases | 1244 | Missing cases | 0 | | |

**APPENDIX 3-4**

A COPY OF THE GAUSS PROGRAM USED TO GENERATE DATA FOR THE GROUP PREDICTOR

---

```
@= THIS IS FOR PROGRAM USED TO CREATING SCHOOL LEVEL VARIABLE
FOR 60 SCHOOL:j=60,i=10,N=600 07/11/91 ============@
new;
output file=c:\gauss\datakam\schvar.dat reset;
j=60;n=600;
i=10;
/*creating ZJ'S */
Z1J=ones(j,1);
Z2J=rndn(j,1);   /* Z2j~N(0,1)  */
Z3 = Z1j~Z2j;
format /rd 8,5;
print Z3 ;
output off;
end;
```

## APPENDIX 3-5

### A COPY OF THE GAUSS PROGRAM USED TO GENERATE DATA FOR THE WITHIN-GROUP PREDICTOR

```
@= THIS IS PROGRAM USED TO CREATE 10 STUDENT WITHIN EACH
SCHOOL FOR 60 SCHOOL J=60,I=10,N=600  07/11/91 ============@
new;
output file=c:\gauss\datakam\STVAR10W.DAT reset;
j=60;n=600;
i=10;
/* generating X's for 600 student in 60 school */
  SD=1;
  DO WHILE SD <= J;
   X1= rndn(I,1);       /*X1~N(0,1)  */
   SD = SD +1;
 @== MX1=meanc(x1);stx1=stdc(x1);vax1=vcx(x1);==@
   FORMAT /RD 8,5 ;
   PRINT X1;;
 ENDO;
end;
```

## APPENDIX 3-6

### A COPY OF THE GAUSS PROGRAM USED TO GENERATE THE DATA SET FOR THE CELL (ND,n10,RRSS)

```
@= j=60, i=10, N=600 ============@
New ,20000;
j=60;N=600;
i=10;
/*creating B1J'S AND B2J'S FOR 60 SCHOOLS*/
@======== gm10= -1.80, gm11= -1.20, gm20= -.5, gm21= .75 , z1j
a vector 60 * 1 of one's z2j a vector 60 * 1 of normal
distribution with mean of 0 and variance of 1  =========@

/* creating  gamma */
gm10= -1.80; gm11= -1.20;
gm20= -.50; gm21= .75;

  load X[N,1] = \gauss\datakam\stVAR10W.dat;
  load Z[J,2] = \gauss\datakam\schvar.dat;
  Z1j = SUBMAT (Z,0,1);
  Z2j = SUBMAT (Z,0,2);
  H = reshape (X,j,i);

/* generating a1j and a2j */

f1="c:\\gauss\\datakam\\studinf.cel";
output file =^f1 reset ;

rr=1;

do while rr <= 100 ;
taj0=rndn(j,1); a1j=0.9975*taj0;     /*a1j~N(0,.995)  */
taj1=rndn(j,1); a2j=0.07071*taj1;    /*a2j~N(0,.005)  */

if rr ==1;
f2="c:\\gauss\\datakam\\spc_01dt.001" ;
f3="c:\\gauss\\datakam\\vac_01dt.001" ;
elseif rr ==2;
f2="c:\\gauss\\datakam\\spc_01dt.002" ;
f3="c:\\gauss\\datakam\\vac_01dt.002" ;
elseif rr ==3;
f2="c:\\gauss\\datakam\\spc_01dt.003" ;
f3="c:\\gauss\\datakam\\vac_01dt.003" ;
elseif rr ==4;
f2="c:\\gauss\\datakam\\spc_01dt.004" ;
f3="c:\\gauss\\datakam\\vac_01dt.004" ;
elseif rr ==5;
```

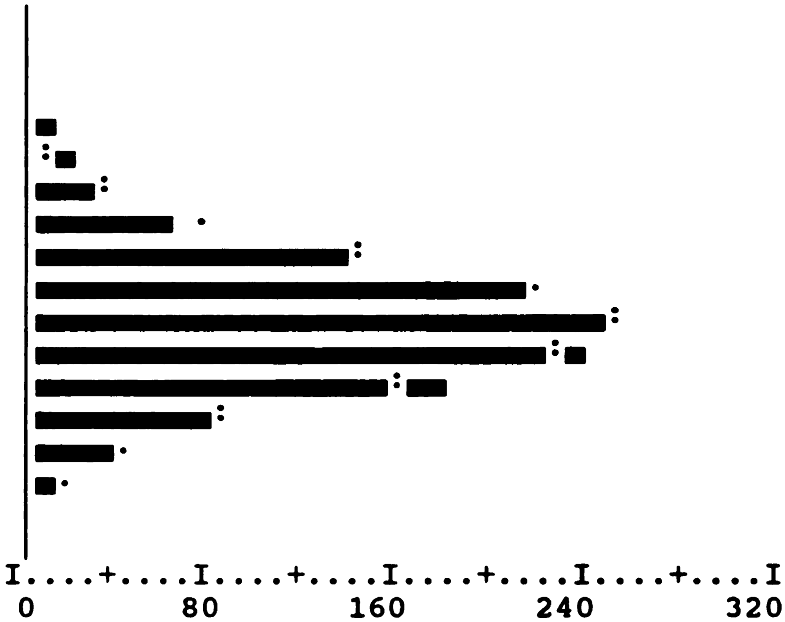
```
  .          .          .
  .          .          .
  .          .          .
f2="c:\\gauss\\datakam\\spc_01dt.095" ;
f3="c:\\gauss\\datakam\\vac_01dt.095" ;
elseif rr ==96;
f2="c:\\gauss\\datakam\\spc_01dt.096" ;
f3="c:\\gauss\\datakam\\vac_01dt.096" ;
elseif rr ==97;
f2="c:\\gauss\\datakam\\spc_01dt.097" ;
f3="c:\\gauss\\datakam\\vac_01dt.097" ;
elseif rr==98;
f2="c:\\gauss\\datakam\\spc_01dt.098" ;
f3="c:\\gauss\\datakam\\vac_01dt.098" ;
elseif rr==99;
f2="c:\\gauss\\datakam\\spc_01dt.099" ;
f3="c:\\gauss\\datakam\\vac_01dt.099" ;
elseif rr==100;
f2="c:\\gauss\\datakam\\spc_01dt.100" ;
f3="c:\\gauss\\datakam\\vac_01dt.100" ;
endif;

/* B1j equations , B2j equations */
B1j = Z1j * gm10 + Z2j * gm11 + a1j ;
B2j = Z1j * gm20 + Z2j * gm21 + a2j ;

L = B1J~B2J;
OUTPUT OFF;
output file =^f1 ;
OUTPUT ON ;
Mb1j=meanc(b1j);stb1j=stdc(b1j);VAB1J=VCX(B1J);
Mb2j=meanc(b2j);stb2j=stdc(b2j);VAB2J=VCX(B2J);
Ma1j=meanc(a1j);sta1j=stdc(a1j);VAA1J=VCX(A1J);
Ma2j=meanc(a2j);sta2j=stdc(a2j);VAA2J=VCX(A2J);
A12=A1J~A2J;
COVA12=VCX(A12);
COVB12=VCX(L);

     Format /rd 8,5;
          PRINT;
          print "**********   The run #   **********";
          PRINT RR;
          PRINT;
          print "Mean stand division variance of B1j";
          print  Mb1j~stb1j~VAB1J;
          print;
          print "Mean stand division variance of B2j";
          print Mb2j~stb2j~VAB2J;
          print;
          print "Mean stand division variance of a1j";
          print Ma1j~sta1J~VAA1J;
          print;
          print "Mean stand division variance of a2j";
```

```
            print Ma2j~sta2j~VAA2J;
            PRINT;
            print "variance covariance matrix of a1j and a2j";
            PRINT COVA12;
            print;
            print "variance covariance matrix of b1j and b2j";
            PRINT COVB12;
            print ;
            print  "-----------------------";
            print ;
            print "-------------------------";

output off;
output file =^f2 reset;
output file =^f3 reset;

K = ones(i,1);

/* generating dependent variable */

   SD=1;
     DO WHILE SD <= j ;
         H1 = submat (H,SD,0);        /*X1~N(0,1)   */
         X1 = H1';
         Y = K * B1j[SD,1] + X1 * B2J[SD,1] ;
         EY1 = exp(Y);
         EY2 = (EY1 + 1 );
         EY  = EY1./EY2;
         u=rndu(i,1);
         dep = ( u .<= ey );
         schvar = K * Z2j[SD,1];
         INTX1Z2=X1 .* SCHVAR ;
         SD  = SD  + 1;

       format /rd 8,5;
       print dep~X1~INTX1Z2;;
       output off;

       output file=^f2;
       output on;
       format /rd 8,5;
       print dep~x1~INTX1Z2~schvar;;
       output off;

       output file=^f3;
       output on ;
    ENDO;   /* end the loop  creating data for each school */
   rr = rr +1   ;
 endo;
/* the end of 100 replication */
end;
```

# APPENDIX 4-1

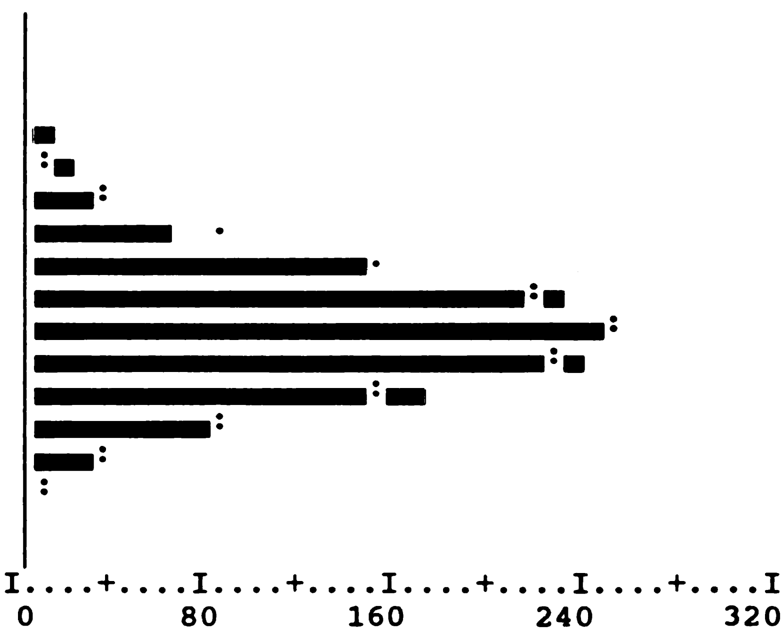
HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{00}$ BY THE SPSS ESTIMATION METHOD.

| Count | Midpoint |
|---|---|
| 1 | -2.250 |
| 0 | -2.175 |
| 2 | -2.100 |
| 4 | -2.025 |
| 16 | -1.950 |
| 29 | -1.875 |
| 54 | -1.800 |
| 136 | -1.725 |
| 203 | -1.650 |
| 236 | -1.575 |
| 225 | -1.500 |
| 171 | -1.425 |
| 80 | -1.350 |
| 33 | -1.275 |
| 8 | -1.200 |
| 2 | -1.125 |
| 0 | -1.050 |




HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{00}$ BY THE VARCL ESTIMATION METHOD.

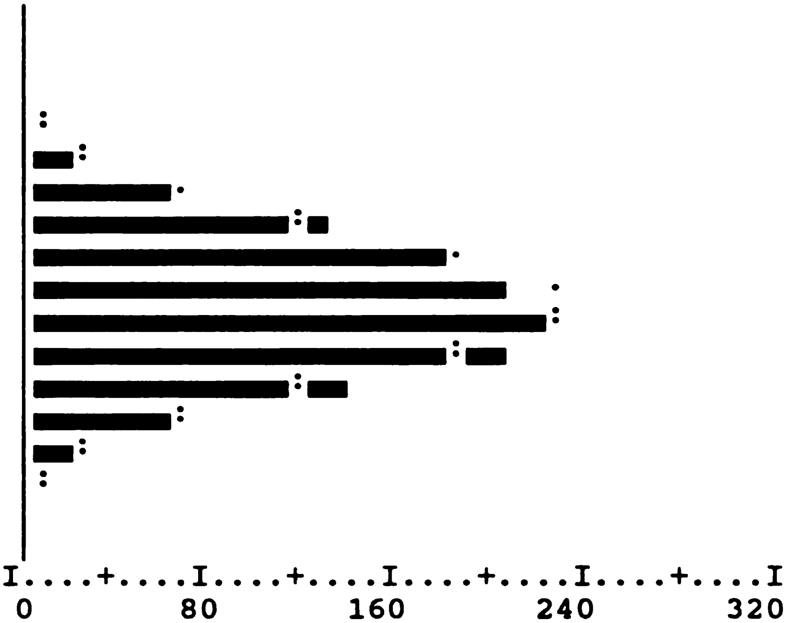
| Count | Midpoint |
|---|---|
| 0 | -2.250 |
| 1 | -2.175 |
| 1 | -2.100 |
| 6 | -2.025 |
| 14 | -1.950 |
| 32 | -1.875 |
| 53 | -1.800 |
| 137 | -1.725 |
| 212 | -1.650 |
| 243 | -1.575 |
| 220 | -1.500 |
| 161 | -1.425 |
| 78 | -1.350 |
| 31 | -1.275 |
| 9 | -1.200 |
| 2 | -1.125 |
| 0 | -1.050 |



# APPENDIX 4-2

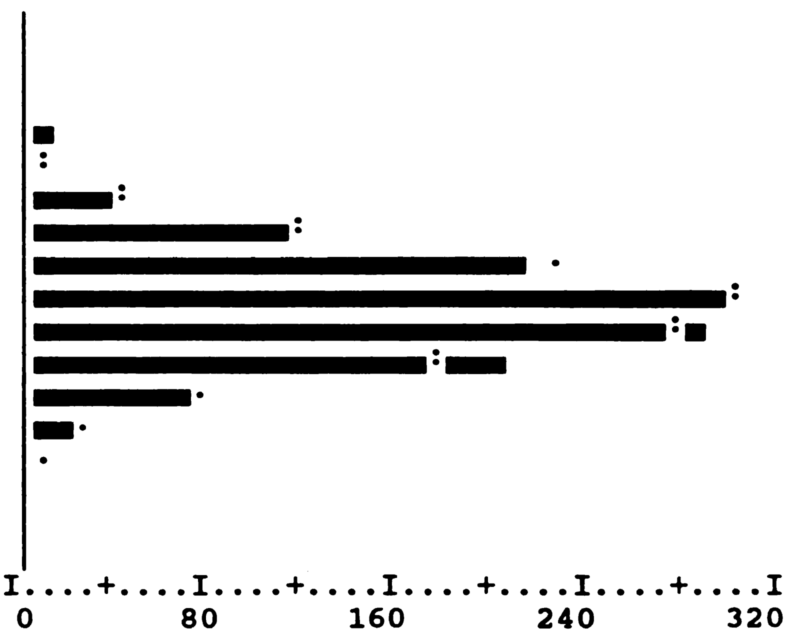
## HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{01}$ BY THE SPSS ESTIMATION METHOD

| Count | Midpoint |
|---|---|
| 2 | -1.700 |
| 3 | -1.625 |
| 3 | -1.550 |
| 11 | -1.475 |
| 22 | -1.400 |
| 54 | -1.325 |
| 121 | -1.250 |
| 165 | -1.175 |
| 195 | -1.100 |
| 217 | -1.025 |
| 192 | -.950 |
| 126 | -.875 |
| 60 | -.800 |
| 21 | -.725 |
| 4 | -.650 |
| 2 | -.575 |
| 2 | -.500 |




## HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{01}$ BY THE VARCL ESTIMATION METHOD

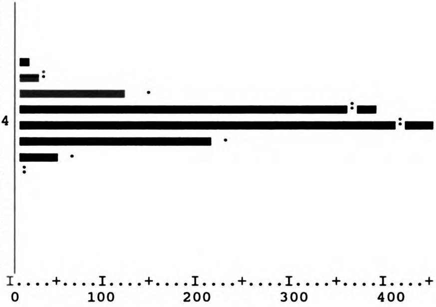
| Count | Midpoint |
|---|---|
| 0 | -1.9 |
| 0 | -1.8 |
| 3 | -1.7 |
| 5 | -1.6 |
| 9 | -1.5 |
| 36 | -1.4 |
| 115 | -1.3 |
| 202 | -1.2 |
| 287 | -1.1 |
| 268 | -1.0 |
| 188 | -.9 |
| 64 | -.8 |
| 18 | -.7 |
| 3 | -.6 |
| 2 | -.5 |
| 0 | -.4 |
| 0 | -.3 |



## APPENDIX 4-3

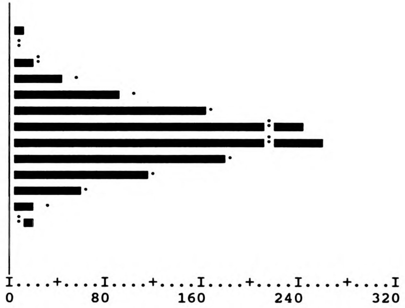
HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{1o}$ BY THE SPSS ESTIMATION METHOD

| Count | Midpoint |
|---|---|
| 0 | -1.1 |
| 0 | -1.0 |
| 1 | -.9 |
| 7 | -.8 |
| 26 | -.7 |
| 107 | -.6 |
| 370 | -.5 |
| 434 | -.4 |
| 201 | -.3 |
| 40 | -.2 |
| 11 | -.1 |
| 2 | .0 |
| 0 | .1 |
| 0 | .2 |
| 0 | .3 |
| 1 | .4 |
| 0 | .5 |



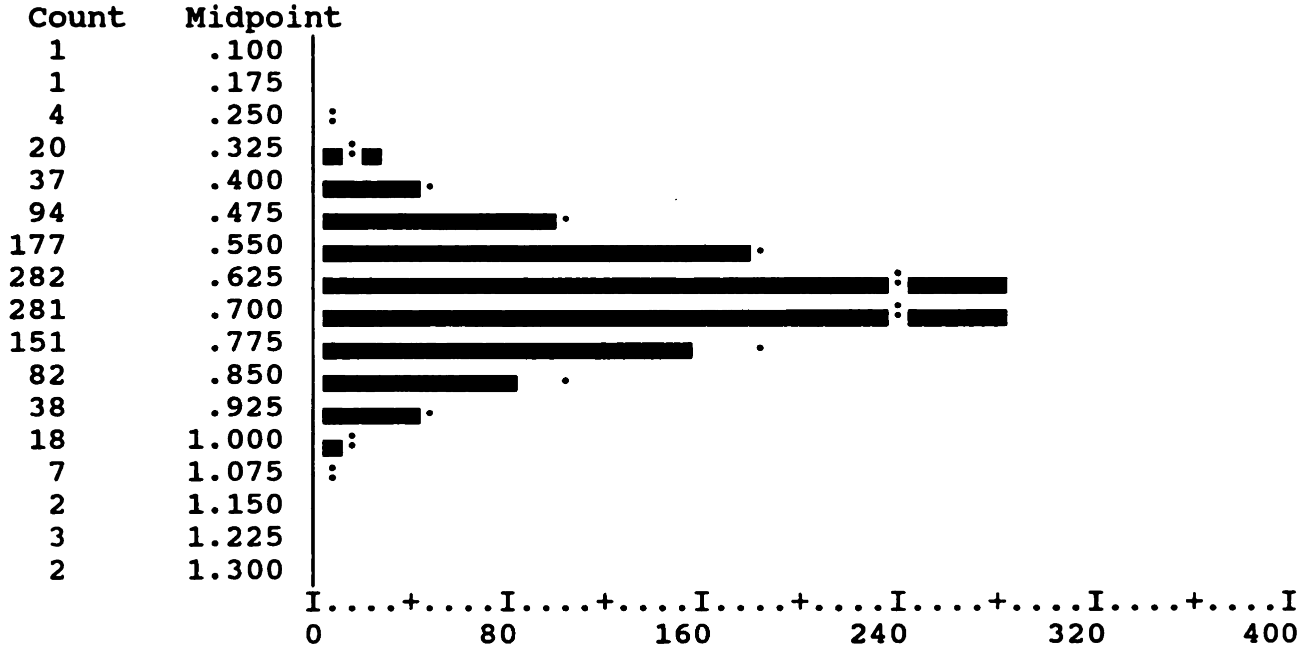

HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{1o}$ BY THE VARCL ESTIMATION METHOD

| Count | Midpoint |
|---|---|
| 1 | -.81 |
| 8 | -.76 |
| 6 | -.71 |
| 24 | -.66 |
| 38 | -.61 |
| 86 | -.56 |
| 159 | -.51 |
| 239 | -.46 |
| 258 | -.41 |
| 172 | -.36 |
| 111 | -.31 |
| 58 | -.26 |
| 18 | -.21 |
| 16 | -.16 |
| 2 | -.11 |
| 3 | -.06 |
| 1 | -.01 |

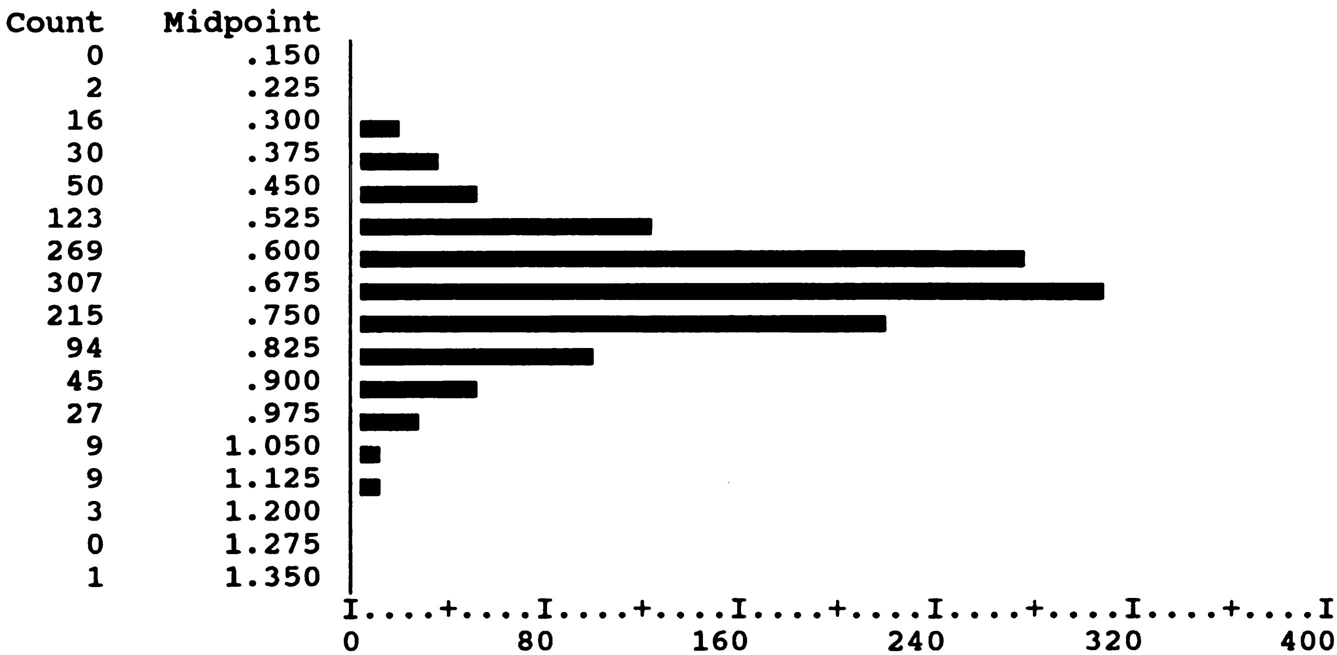


## APPENDIX 4-4

HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{11}$ BY THE SPSS ESTIMATION METHOD




HISTOGRAM FREQUENCY FOR ESTIMATED $\gamma_{11}$ BY THE VARCL ESTIMATION METHOD



# APPENDIX 5-1

## THE FIRST AND SECOND DERIVATIVE OF THE LOGISTIC FUNCTION

The logistic function

$$f(\alpha) = \frac{e^{\alpha}}{1+ e^{\alpha}} = \frac{1}{1+ e^{-\alpha}}$$

The first derivative of the logistic function

$$f'(\alpha) = \frac{(-1)(-e^{-\alpha})}{(1+ e^{-\alpha})^2} = \frac{e^{-\alpha}}{(1+ e^{-\alpha})^2}$$

The second derivative of the logistic function

$$f''(\alpha) = \frac{(1+ e^{-\alpha})^2(-e^{-\alpha}) \;-\; e^{-\alpha}\, 2(1+ e^{-\alpha})(-e^{-\alpha})}{(1+ e^{-\alpha})^4}$$

$$= \frac{(1 + 2e^{-\alpha} + e^{-2\alpha})(-e^{-\alpha}) + e^{-2\alpha}\, 2(1 + e^{-\alpha})}{(1+ e^{-\alpha})^4}$$

$$= \frac{-e^{-\alpha} -2e^{-2\alpha} -e^{-3\alpha} + 2e^{-2\alpha} +2e^{-3\alpha}}{(1+ e^{-\alpha})^4}$$

$$= \frac{-e^{-\alpha} + e^{-3\alpha}}{(1+ e^{-\alpha})^4}$$

$$= \frac{e^{-\alpha}(e^{-2\alpha} -1)}{(1+ e^{-\alpha})^4} > 0 \quad if \quad \alpha < 0$$

# BIBLIOGRAPHY

BIBLIOGRAPHY

Aitkin, M., Anderson, D., and Hinde, J. (1981). Statistical modeling of data on teaching styles. Journal of the Royal statistical Society, Series A, 144 (4), 419-461.

Aitkin, M., and Longford, N. (1986). Statistical modelling issues in school effectiveness Studies (with discussion). Journal of the Royal statistical Society, Series A, 149 (1), 1-43.

Alba, R. D. (1987). Interpreting Parameters of Log-Linear Models. Sociological Methods & Research, 16(1), 45-77.

Alker, H. R. (1969). A Topology of Ecological Fallacies, in M. Dogan and S. Rokkan (Eds.), Social Ecology. Cambridge, MA: MIT Press.

Anderson, J. A. (1980). Logistic Discrimination, in P. R. Krishnaiah and L. Kanal (Eds.), Handbook of Statistics. Vol. 2: Classification Pattern Recognition, and Reduction of Dimension, New York: North-Holland.

Anderson, D. A., and Aitkin, M. (1985). Variance Component Models with Binary Response: Interviewer Variability. Journal of the Royal Statistical Society, Series B, 47(2), 203-210.

Bock, R. D. (1983). The Discrete Bayesian. In H. Wainer and Messick (eds.), Principles of Modern Psychological Measurement: A Festschrift for Fredrick M. Lord. Hillsdale, N.J.: Lawrence Erlebaum Associates.

Box, G. E. P., and Tiao, G. C. (1973). Bayesian Inference in Statistical Analysis. Reading, MA: Addison-Wesley.

Braun H. (1989). Empirical Bayes Methods: A Tool for Exploratory Analysis. In Multilevel Analysis of Educational Data, edited by D. Bock. New York: Academic Press.

Breslow, N. E., and Day, N. E. (1980). Statistical Methods in Cancer Research: Vol 1, The Analysis of Case-Control Studies. Lyon, France: International Agency on Cancer.

Bryk, A. S., Raudenbush, S. W., Seltzer, M., and Congdon (1988). An Introduction to HLM: Computer Program and User's Guide. Version 1.0, University of Chicago.

Bull, S. B., and Donner, A. (1987). The Efficiency of Multinomial Logistic Regression Compared with Multiple Group Discriminant Analysis. Journal of the American Statistical Association, 82, 1118-1122.

Burstein, L. (1980). The Analysis of Multilevel Data in Educational Data in Educational Research and Evaluation. Review of Research in Education, 8, 158-233.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. Review of Economic Studies, 47, 225-238.

Clogg, C. C., Rubin, D. B., Schenker, N., Schultz, B, and Weidman, L. (1990). Multiple Imputation of Industry and Occupation Codes in Census Public-Use Samples Using Bayesian Logistic Regression. (In Press).

Cox, D. R. (1970). The Analysis of Binary Data. London: Chapman and Hall.

Cox, D. R. & Snell, E. J. (1989). The Analysis of Binary Data. Second edition. London: Chapman and Hall.
Dempster, A.P.; Laird, N.M.; and Rubin, D.B. (1977). Maximum Likelihood From Incomplete Data Via the EM Algorithm (with discussion. Journal of the Royal Statistical Society, Series B, 39, 1-8.

Dempster, A. P., D. Rubin, & R. Tsutakauwa (1981). Estimation in Covariant Component Models. Journal of American Statistical Association, 76: 341-353.

Draper, N. and Smith, H. (1966). Applied Regression Analysis. New York: Wiley.

Dunteman, G. H. (1984). Introduction to Linear Models. Beverly Hills, California: Sage Publications.

Efron, B. (1975). The Efficiency of Logistic Regression Compared to Normal Discriminant Analysis. Journal of the American Statistical Association, 70, 892-898.

Geisser, S. (1984). On Prior Distributions for Binary Trials. The American Statistician, 38, 244-247.

Glass, G. V., and Smith, M. L. (1979). Meta-analysis of the Research on Class Size and Achievement. Educational Evaluation and Policy Analysis, 1, 2-16.

Goldstein, H. I. (1986). Multilevel Mixed Linear Model analysis Using Iterative Generalized Least Squares. Biometrika, 73(1), 43-56.

Goldstein, H. I. (1987). Multilevel Models in Educational and Social Research. London: Oxford University Press.

Goldstein, H. I. (1990). Nonlinear Multilevel models, with an Application to Discrete Response data. (Submitted for publication).

Haberman, S. J. (1974). The Analysis of Frequency Data. Chicago: University of Chicago Press.

Haberman, S. J. (1977). Log-Linear Models and Frequency Tables With Small Expected Counts. Annals of Statistics, 5, 1148- 1169.

Haney, W. (1980). Units and Levels of Analysis in Large Scale Evaluation. New Directions for Methodology of Social and Behavioral Sciences, 6, 1-15.

Hannan, M.T. (1971). Aggregation and Disaggregation in Sociology. Lexington, MA: D.C. Heath.

Hanushek, E. A., and Jackson, J. E. (1977). Statistical Methods for Social Scientists. New York: Academic Press.

Hosmer, D. W., and Lemeshow S. (1989). Applied Logistic Regression. New York: Wiley & Sons.

Kreft, I. G. (1987). Methods and Models for the Measurement of School Effects. Dissertation, University of Amsterdam.

Kreft, I. G., and Kim, K. (1990a). GENMOD, HLM, ML2 and VARCL, Four Statistical Packages for Hierarchical Linear Regression. In Theory and Model in Multilevel Research: Convergence or Divergence ? edited by Eeden, P., Hox, J., and Hauer, J. Amsterdam:SISWO.

Kreft, I. G., De Leeuw, J. , and Kim, K. S. (1990b). Comparing Four Different Packages for Hierarchical Linear Regression GENMOD,HLM, ML2 and VARCL", CSE Report # 310, Center for the Study of Evaluation, UCLA.

Korn, E. L., and Whittemore, A. S. (1979). Methods for Analyzing Panel Studies of Acute Health Effects of Air Pollution. Biometrics, 35, 795-804.

Laird, N. M. (1978). Empirical Bayes Methods for Two-Day Contingency Tables. Biometrika, 65, 581-590.

Laird, N. M., and Louis, T. A. (1982). Approximate Posterior Distributions for Incomplete Data Problems. Journal of the Royal Statistical Society, Series B, 44, 190-200.

Langbein, L. I. (1977). Schools or Students: Aggregation Problems in the Study of Student Achievement. Evaluation Studies Review Annual, 2, 270-298.

Leonard, T. (1972a). Bayesian Methods for Binomial Data. Biometrika, 59(3), 581.

Leonard, T. (1972b). Bayesian Methods for Discrete Data. Act Technical Bulletin, No.10, 1-20.

Leonard, T. (1975). A Bayesian Approach to the Linear Model with Unequal Variances. Technometrics, 17(1), 95-102.

Lindley, D. V., and Smith, A.F.M. (1972). Bayes Estimates for the Linear Model. Journal of the Royal Statistical Society, Series B, 34, 1-41.

Longford N. T. (1992). Private communication.

Longford N. T. (1988). A Quasilikelihood Adaptation for Variance Component Analysis. Presented at the Annual Meeting of ASA, New Orleans, LA, August 1988.

Longford N. T. (1990). Multilevel Modelling into the 1990S. Multilevel Modelling Newsletter, 2(1), 2.

Mason, W.M.; Wong, G.Y.; and Entwisle, B. (1984). Contextual Analysis Through the Multi-level Linear Model. Sociological Methodology. San Francisco, CA: Jossey Bass, 72-103.

McCullagh, P. (1983). Quasi-likelihood functions. Ann. Statist. 11, 59-67.

McCullagh, P. , and Nelder J. A. (1983). Generalized Linear Models. London: Chapman and Hall.

Nelder, J. A., and Pregibon, D. (1987). An Extended Quasilikelihood Function. Biometrika, 74, 221-232.

Prosser R., Rasbash J., and Goldstein H. (1991). ML3 Software for Three-level Analysis Users' Guide for V.2, Institute of Education, University of London.

Randenbush, S.W. (1988). Educational Applications of Hierarchical Linear Models: A Review", Journal of Educational statistics, 13 (2), 85-116.

Randenbush, S. W., and C. Bhumirat (1989). Results of a National Survey of Primary Education in Thailand, Part I: School Size, Sector, Facilities, Equipment, Textbook and Teaching Materials as Predictors of Pupil Achievement.

Randenbush, S. W., and Bryk, A. S. (1986). A Hierarchical Model for Studying School Effects. Sociology of Education, 59, 1-17.

Randenbush, S. W., & A. S. Bryk (1988). Methodological Advances in Analyzing the Effects of Schools and Classrooms on Student Learning. Review of Research in Educations, 15, 423-476.

Robinson, W.S. (1950). Ecological Correlations and the Behavior of Individual. American Sociological Review, 15, 351-357.

Rubin, D. B., and Schenker, N. (1987). Logit-Based Interval Estimation for Binomial Data Using the Jeffreys Prior. pp. 131-144 in C.C. Clogg, ed., Sociological Methodology 1987, Washington, D. C.: American Sociological Association.

Scheffe, H. (1959). The Analysis of Variance. New York: Wiley.

Shigemasu, K. (1976). Development and Validation of a Simplified m-Group Regression Model. Journal of Educational statistics, 1(2), 157-180.

Stiratelli, R., Laird, N., and Ware, J. H. (1984). Random Effects Models for Serial Observations With Binary Response. Biometrics, 40, 961-971.

Walsh, J. D. (1947). Concerning the Effect of the Intraclass Correlation on Certain Significance Test. Annals of Mathematical Statistics, 18, 88-96.

Wedderburn, R. W. M. (1974). Quasi-likelihood functions, generalized linear models and the Gauss-newton method. Biometrika, 61, 439-47.

Weisberg S. (1985). Applied Linear Regression. New York: Wiley & Sons.

Wong, G.Y., and Mason, W.M. (1985). The Hierarchical Logistic Regression Model for Multilevel Analysis. Journal of American Statistical Association, 80, 513-524.