# Florida State University Libraries

2011

# Meta-Analytic Structural Equation Modeling (MASEM): Comparison of the Multivariate Methods

Ying Zhang

THE FLORIDA STATE UNIVERSITY

COLLEGE OF EDUCATION

META-ANALYTIC STRUCTURAL EQUATION MODELING (MASEM):

COMPARISON OF THE MULTIVARIATE METHODS

By

YING ZHANG

A Dissertation submitted to the
Department of Educational Psychology and Learning Systems

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Degree Awarded:
Spring Semester, 2011

The members of the committee approve the dissertation of Ying Zhang defended on Jan. 14, 2011.

_____
Betsy Jane Becker
Professor Directing Dissertation

_____
Yanyun Yang
Professor Co-Directing Dissertation

_____
Xufeng Niu
University Representative

_____
Robert Eklund
Committee Member

Approved:

_____
Betsy Jane Becker, Chair, Department of Educational Psychology and Learning Systems

The Graduate School has verified and approved the above-named committee members.

For my family

# ACKNOWLEDGEMENTS

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

Meta-analytic Structural Equation Modeling (MASEM) has drawn interest from many researchers recently. In doing MASEM, researchers usually first synthesize correlation matrices across studies using meta-analysis techniques and then analyze the pooled correlation matrix using structural equation modeling techniques. Several multivariate methods of MASEM have been proposed by the researchers. In this dissertation, I compared the commonly used multivariate methods for meta-analytic path modeling. Specifically, I examined the Generalized Least Squares (GLS) method (Becker, 1992; Becker & Schram, 1994) and the Two-Stage Structural Equation Modeling (TSSEM) method (Cheung, 2002; Cheung & Chan, 2005) using both simulation studies and real data analyses. Both the traditional GLS approach (Becker, 1992) and the modified GLS approaches (Becker & Fahrbach, 1994) were applied and compared with the TSSEM approach. Fixed-effects data and random-effects data were generated to see how these approaches differ at the first and second stages of MASEM. The results shows that the modified GLS approach performs as well as or better than the TSSEM approach in both the first step of synthesizing correlation matrices and the second step estimation of the parameters and standard errors, using both fixed-effects data and random-effects data. The original GLS approach only performs well when the within-study sample size is large enough (of the simulation situations in this dissertation, $n \geq 100$). Both the modified GLS approach and the TSSEM approach produce equivalent parameter estimates across all conditions. However, the standard errors from the TSSEM approach seem to be over-estimates under certain conditions. Overall, both the modified GLS and TSSEM approaches are appropriate for conducting meta-analytic path modeling and the difference in parameter estimates is minimal.

# CHAPTER I

# INTRODUCTION

Meta-analysis is a set of statistical techniques to examine results from a series of related studies. Glass (1976) coined the term "meta-analysis" to refer to the "analysis of analyses."(p. 3). It is widely used to synthesize research findings in educational, psychological, behavioral and medical sciences (e.g., Hedges & Olkin, 1985; Hunter & Schmidt, 2004; Shadish, 1996; Sutton et al., 2000). It uses effect sizes as quantitative indices to summarize the results of studies. Common effect sizes include Pearson's correlation coefficient $r$, standardized mean differences (Cohen's $d$ and Hedges's $g$) and odds ratios. Generally, researchers use meta-analysis to test homogeneity among the effect sizes across studies. If the effect sizes are homogeneous, the reviewer may estimate the mean effect size, construct a confidence interval for the mean effect size and test hypotheses about it. If the effect sizes are found to be heterogeneous, the reviewer may apply moderator analyses to explain the variations of effect sizes.

## Methods for Meta-analytic Structural Equation Modeling

Several textbooks have given comprehensive descriptions of the concepts and techniques of meta-analysis (e.g., Cooper & Hedges, 1994; Hedges & Olkin, 1985; Hunter & Schmidt, 1990, 2004; Lipsey & Wilson, 2001; Sutton et al., 2000). Most meta-analyses involve a single effect size like the standardized mean difference between an experimental group and a control group, or the Pearson's correlation coefficient $r$ between two outcomes; recently, growing interest has been shown in the synthesis of correlation matrices across studies. Similar to the research synthesis of a single effect size per study, the synthesis of correlation matrices involves firstly testing the homogeneity of correlation matrices across studies; a pooled correlation matrix is then obtained if homogeneity is found. If correlation matrices are heterogeneous, moderator analysis may

be applied to explain the variations. Methodological approaches have been proposed and developed for synthesizing correlation matrices and conducting subsequent analyses (e.g., methods to estimate a regression or path model, or compute structural equation models) using the pooled correlation matrix (e.g., Becker, 1992, 1995, 2000; Becker & Schram, 1994; Beretvas & Furlow, 2006; Cheung & Chan, 2005; Hafdahl, 2001, 2007, 2009; Shadish, 1996; Viswesvaran & Ones, 1995).

As noted by Shadish (1996), meta-analysis has long been focused on describing single effects, and this practice had "severely limited its capacity to contribute to one of the most fundamental tasks of science --- the development of explanatory theories." (p. 48). He suggested that researchers should conduct 'causal mediating modeling' (which uses the path analysis techniques and structural equation modeling (SEM) techniques) to address this issue.

Cheung and Chan (2005) used the term "Meta-analytic Structural Equation Modeling (MASEM)" as a general description for these techniques. They suggested that more specific techniques, like meta-analytic path analysis which does not consider latent variables, and meta-analytic confirmatory factor analysis which considers only the measurement model can be considered as special cases of MASEM. Other researchers have used similar terminologies to describe similar procedures, for instance, meta-analytic path analysis (Colquitt, LePine, & Noe, 2000), meta-analysis of factor analysis (G. Becker, 1996).

Generally there are two steps involved in MASEM: the first step is the synthesis of correlation coefficients across studies; the second step is to apply structural equation modeling techniques to explore the relationship among variables using the pooled correlation matrix (Viswesvaran & Ones, 1995). There are several different methods to choose from in either stage of MASEM. At the first step, univariate and multivariate methods have been proposed for synthesizing the correlation coefficients. The univariate methods proposed by Hedges and Olkin (1995) weight each effect size by its sampling variance and then synthesize the weighted effect sizes. Hedges and Olkin (1985) also suggested using Fisher's $z$ transformation to normalize and stabilize the variance of the correlation coefficients. Hunter and Schmidt (1990, 2004) proposed to meta-analyze the correlation coefficients using a sample-size-weighting scheme. These univariate methods

ignore the covariation among the correlation coefficients, thus may lead to inaccurate estimation of parameters and standard errors.

Multivariate methods of synthesizing correlation matrices have also been proposed by several researchers (e.g., Becker, 1992, 1995; Becker & Fahrbach, 1994; Cheung & Chan, 2005; Hafdahl, 2001, 2007). Unlike the univariate methods, the multivariate methods consider the dependence of the correlation coefficients in the correlation matrices.

At the second stage of MASEM, Becker (1992, 1995) proposed a direct computation method based on Generalized Least Squares (GLS) for linear models and path analyses using the synthesized correlation matrix and its asymptotic covariance matrix. Cheung (2002) used common SEM software packages (i.e., LISREL, Jöreskog & Sörbom, 1999) to fit path analytic models and confirmatory factor analysis models using the pooled correlation matrix and its asymptotic covariance matrix from the first step. Cheung and Chan (2005) called this method the 'Two-stage Structural Equation Modeling' (TSSEM) approach.

Several studies have addressed other important issues in MASEM, for example, missing data and analyzing correlation matrices rather than covariance matrices using structural equation models (e.g., Furlow & Beretvas, 2005; Beretvas & Furlow, 2006; Cheung & Chan, 2009).

To date, only a few studies have examined the performance of the aforementioned methods in the second stage of MASEM (e.g., Cheung, 2002, 2005; Furlow & Beretvas, 2005). No study to date has examined the performance of Becker's (1992, 1995) direct computation method as opposed to the meta-analytic path-analysis method proposed by Cheung (2002).

**Fixed-Effects Model vs. Random-Effects Model**

Hedges and Vevea (1998) described the distinction between fixed- and random-effects models. In the fixed-effects models, all the effect sizes are assumed to share an unknown but constant population effect size, thus there is one common population parameter for all effect sizes in the meta-analysis, thus we say that all the study effect

sizes are homogeneous. On the other hand, random-effects models assume a distribution of the population effect sizes, thus the population effect sizes differs randomly from study to study. The population effect sizes can be thought of as a sample from a universe of possible effects, that is, a 'super/hyper population' (Becker, 1996; Hedges, 1992). Then, the studies are heterogeneous in this situation. Usually, random-effects models are thought to be more realistic in practice, especially when the researchers want to generalize the results to a broad domain rather than the studies in hand (see Hedges & Vevea, 1998). Statistically, the main difference between these two types of models is the error variance associated with the synthesized effect size: the error term in a fixed-effects model contains only within-study variations, while the error term in a random-effects model contains both within- and between-study variations.

The TSSEM (Cheung & Chan, 2005) approach assumed a fixed-effects model for the synthesis of correlation matrices and provided no obvious means of modeling random-effects models. Becker (1992, 1995) also described a method for the random-effects model using a GLS approach. However, this GLS method for the random-effects model has not been further and subsequently studied. To the best of my knowledge, how random-effects models affect parameter estimation at the second stage of MASEM has not been explored using either of the multivariate methods described above.

<div align="center">

**Purpose of the Dissertation**

</div>

The first main purpose of this dissertation is to compare Becker's GLS approach (1992, 1995) and Cheung and Chan's (2005) TSSEM approach for meta-analytic path analyses. Specifically, the GLS methods and TSSEM methods will be used in synthesizing the correlation matrices. The resulting pooled correlation matrices are then subject to the path modeling with Asymptotic Distribution-free (ADF) estimation as the estimation method, and path modeling using Becker's (1992) direct computation method. Given the fact that a random-effects model is more appropriate in many situations in practice, the second purpose of this dissertation is to explore the method of estimating path coefficients based on an average correlation matrix from the random-effects model

and compare the parameter estimates obtained using the GLS method and TSSEM method.

## Outline of the Chapters

In Chapter II, I will introduce existing methods for synthesizing correlation matrices, especially the two multivariate methods: GLS approach and TSSEM approach. Path modeling processes using each of these two multivariate approaches are presented in Chapter III. In Chapter IV, I report several simulation studies to compare the parameter estimates of GLS and TSSEM approach under both fixed-effects models and random-effects models. A real example of data analysis is presented in Chapter V to show how these methods are applied using real data. Finally, discussions and future research directions are presented in Chapter VI.

# CHAPTER II

# LITERATURE REVIEW

Meta-analytic Structural Equation Modeling (MASEM) techniques have been increasingly applied in research fields in psychological, educational and behavioral sciences. For example, Colquitt, Scott, and LePine (2007) used a meta-analytic path analysis to examine the relationship between trust variables, risk taking, and job performance. Whiteside and Becker (2000) applied a model-driven meta-analysis to examine the relationship among parental factors and child's postdivorce adjustment. Brown et al. (2008) applied a meta-analytic path analysis to understand the social cognitive predictors of college students' academic performance and persistence.

Generally, MASEM involves two steps in analyses. At the first step, correlation matrices are synthesized across studies and at the second step, path analysis, factor analysis or structural equation modeling techniques are applied to analyze the pooled correlation matrix. A series of methods have been developed for each of the steps. Univariate and multivariate methods have been proposed to synthesize correlation matrices across studies at the first step. With univariate meta-analysis, the correlation coefficients in the correlation matrix are treated as independent and the correlations among them are ignored (e.g., Schmidt, Hunter & Outerbridge, 1986). Basically, meta-analysis is applied to each element of the correlation matrix and then the synthesized correlation coefficients form a pooled correlation matrix. Probably the most widely used univariate methods are those proposed by Hedges and Olkin (1985) and Hunter and Schmidt (1990, 2004). Both of these univariate approaches provide methods of synthesizing a single correlation coefficient per study, yielding point and interval estimates as well as testing homogeneity of the correlations.

Multivariate methods for synthesizing correlation matrices consider the dependency among the correlation coefficients in the correlation matrix. The variance and covariance matrix among the correlation coefficients is obtained and used in synthesizing the correlation matrices. Later analyses of the pooled correlation matrix

(e.g., exploratory/confirmatory factor analysis, path modeling or linear models) then account for this dependence. Becker (1992, 1995), Becker and Schram (1994), Cheung and Chan (2005), and Furlow and Beretvas (2005) illustrated a range of multivariate techniques to account for the dependence among the elements of the correlation matrices.

In this section, I will summarize available univarate and multivariate methods of synthesizing correlation matrices, especially the multivariate methods proposed by Becker (1992), Becker and Schram (1994), and Cheung and Chan (2005).

## Univariate Methods for Synthesizing Correlation Matrices

Hedges and Olkin (1985) introduced methods for synthesizing bivariate correlation coefficients with fixed-effects and random-effects models. Denote $r_i$ as the observed correlation coefficient in study $i$, $i = 1,...,k$. Assuming a fixed-effects model, the effect size can be expressed as

$$r_i = \rho + e_i$$

where $\rho$ is the population correlation coefficient and $e_i$ is the sampling error in study $i$. Assume that $e_i$ is distributed normally with mean of zero and variance $\sigma_i^2$. The variance $\sigma_i^2$ is the within-study variance and usually assumed known in meta-analysis.

Before averaging the correlation coefficients, a homogeneity test is usually conducted to test if the correlation coefficients are the same in the population. Under the null hypothesis $\rho_1 = \rho_2 = ... = \rho_k$, Hedges and Olkin (1985) showed that the test statistic for testing homogeneity of the correlation coefficients is

$$Q_{HOr} = \sum_{i=1}^{k} w_i (r_i - r.)^2 .$$

where $w_i = 1/\sigma_i^2$ is the weight and $k$ is the total number of studies, and $r.$ is the weighted mean effect size. Under the null hypothesis, the $Q_{HOr}$ statistic has an approximate chi-square distribution with $(k-1)$ degrees of freedom. The weighted average of the effect sizes $r_i$ is calculated as:

$$r. = \frac{\sum_{i=1}^{k} w_i r_i}{\sum_{i=1}^{k} w_i} .$$

The sampling variance of the mean effect size $r.$ is the reciprocal of the sum of the weights in each study, that is,

$$Var(r.) = 1/\sum_{i=1}^{k} w_i .$$

It is known that the sampling distribution of correlation coefficients is skewed when the population correlation coefficients deviates from zero. Hedges and Olkin (1985) proposed to use Fisher's $z$ transformation when averaging the correlation coefficients, where

$$z = \frac{1}{2} \log(\frac{1+r}{1-r}) ,$$

where $z$ represents the transformed raw correlation coefficient $r$, and $\log(.)$ is the natural logarithm. The test for the hypothesis that the transformed correlation coefficients are equal in the population is

$$Q_{HOz} = \sum_{i=1}^{k} (n_i - 3)(z_i - z.)^2 ,$$

where $z_i$ is the transformed correlation coefficients in study $i$, and $z.$ represents the averaged correlation in the z metric. The mean $z.$ can be expressed as

$$z. = \frac{\sum_{i=1}^{k} w_i z_i}{\sum_{i=1}^{k} w_i} ,$$

where $w_i = (n_i - 3)$, with $n_i$ being the sample size in the $i$th study and $k$ the total number of studies. The sampling variance of the $z$-transformed correlations now depends only on the sample size, that is,

$$var(z.) = \frac{1}{(n_i - 3)} .$$

The mean z. can be transformed back to the metric of the correlation coefficient by

$$r. = \frac{\exp(2z.) - 1}{\exp(2z.) + 1}.$$

Hunter and Schmidt (1990, 2004) computed the average correlation coefficient $\bar{r}$ by weighting each correlation coefficient by the study sample size. That is,

$$\bar{r} = \frac{\sum_{i-1}^{k} n_i r_i}{\sum_{i-1}^{k} n_i}$$

where the $n_i$ and $r_i$ are the sample size and sample correlation coefficients in each study as defined above.

The test statistic used by Hunter and Schmidt (1990) is

$$Q_{HSr} = \sum_{i=1}^{k} \frac{(n_i - 1)(r_i - \bar{r})^2}{(1 - \bar{r}^2)^2}.$$

It is asymptotically distributed as a chi-square with ($k$-1) degrees of freedom under the null hypothesis (see Hunter & Schmidt, 1990, pp.110 – 112).

There has long been a controversial view regarding if transformed or untransformed correlation coefficients should be used in meta-analysis (e.g., Hedges & Vevea, 1998; Hunter et al., 1996; Silver and Dunlap, 1987). Field (2001) reported several simulation studies to compare Hunter and Schmidt's correlation method (HS$r$) and Hedges and Olkin's Fisher $r$ to $z$ method (HO$z$) with both homogeneous and heterogeneous data. He found that both methods produce comparable mean estimates under the fixed-effects model, that is, when the correlations are homogeneous. However, he found that the HS$r$ method was too liberal in controlling the Type I error rate for the significance test. He also concluded that for the heterogeneous cases, the HO$z$ method produced biased estimate of the mean effect size, rather the HS$r$ method gave relative accurate estimates. Field's (2000) study was cited by Hunter and Schmidt (2004) as evidence that Fisher-$z$ transformation should not be applied. However, Hafdahl and Williams (2009) revisited the simulation conditions by Field (2001, 2005) and concluded that the "cautions about poor performance of Fisher-z method are largely unfounded, especially with a more appropriate $z$-to-$r$ transformation" (p. 24). An alternative integral $z$-to-$r$ transformation was given by Hafdahl (2009).

Hedges and Olkin (1985, p. 108) pointed out that most procedures available for meta-analysis assume that the outcomes to be analyzed are independent. With univariate methods for synthesizing correlation matrices, covariation among the correlation coefficients is totally ignored. However, these correlations are actually correlated to a certain extent (Olkin & Siotani, 1976). Problems may arise when dependence among the correlation coefficients is ignored or handled inadequately. Becker (2000) pointed out that "one common consequence of ignoring dependence is a possible effect on Type I error level and accuracy of probability statements made for the observed data…… This can be manifest via the values of estimates and standard errors; thus dependence can also affect bias and precision in estimation as well." (p. 503). The better strategy is to consider the dependence among the elements of correlation matrix in synthesizing correlation matrices across studies. To model the dependence among effect sizes, several approaches have been proposed (e.g., Becker, 1992, 1995; Gleser & Olkin, 1994; Hedges & Olkin, 1985; Kalaian & Raudenbush, 1996; Raudenbush, Becker & Kalaian, 1988). These techniques all require that the information of the correlations among the effect sizes be incorporated into the analysis. Specifically, several researchers have proposed or studied multivariate methods for synthesizing correlation matrices (e.g., Becker, 1992, 1995; Beretvas & Furlow, 2006; Cheung & Chan, 2005; Hafdahl, 2001, 2007).

## Multivariate Methods for Synthesizing Correlation Matrices

*Generalized Least squares (GLS) Method*

Becker (1992, 1995) uses generalized least squares (GLS) estimation techniques to model the dependency between correlations when pooling correlation matrices. For each study $i$ with sample size $n_i, i = 1,...,k,$ across $k$ independent studies, there are $p* = p(p-1)/2$ distinct correlation coefficients among $p$ multivariate normally distributed variables. For ease of introduction, we assume that there is no missing data in each study. However, this notation can be extended readily to studies with incomplete data (see Becker, 1992). For example, with $p = 4$ outcomes, we obtain $p* = p(p-1)/2 = 4(4-1)/2 = 6$ distinct correlations for each study. Thus $\mathbf{r}_i, \ i = 1,...,k$ ,

is a column vector which contains these six observed correlations. Denote the $\mathbf{r}_i$ as the estimate of the population correlation vector $\boldsymbol{\rho}_i$. Thus the model

$$\mathbf{r}_i = \boldsymbol{\rho}_i + \mathbf{e}_i \qquad (2.1)$$

expresses the variation of the observed correlations around their population values.

Assume that $\mathbf{e}_i \sim N_{P*}(0, \mathbf{V}_i)$, where $\mathbf{V}_i$ is the variance-covariance matrix of the observed correlations. The dimension of $\mathbf{V}_i$ is $p*\times p*$ or $6\times 6$ here . Olkin and Siotani (1976) derived the large-sample normal approximate variance and covariance matrix for a vector of correlation estimates. For study $i$, the large-sample population variance of the correlation estimate between variable $s$ and $t$ in study $i$, $r_{ist}$ is

$$Var(r_{ist}) = \frac{(1-\rho_{ist}^2)^2}{n_i} , \qquad (2.2)$$

with $\rho_{ist}$ being the population correlation and $n_i$ being the sample size for study $i$. The large-sample covariance, $Cov(r_{ist}, r_{iuv})$ between correlations $r_{ist}$ and $r_{iuv}$ is

$$\begin{aligned} Cov(r_{ist}, r_{iuv}) = [0.5\rho_{ist}\rho_{iuv}(\rho_{isu}^2 + \rho_{isv}^2 + \rho_{itu}^2 + \rho_{itv}^2) + \rho_{isu}\rho_{itv} + \rho_{isv}\rho_{itu} \\ -(\rho_{ist}\rho_{isv}\rho_{isu} + \rho_{its}\rho_{iiu}\rho_{itv} + \rho_{ius}\rho_{iut}\rho_{iuv} + \rho_{ivs}\rho_{ivt}\rho_{ivu})]/n_i \end{aligned} . \qquad (2.3)$$

Since the population parameter, $\rho_{ist}$, is unknown, the standard approach for obtaining the estimates of the variance and covariances for the correlations is to substitute sample estimates, $r_{ist}$, for the corresponding population values and then treat $\mathbf{V}_i$ as known. In theory, the use of GLS for synthesizing correlation matrices should result in more accurate estimates because it accounts for the dependency among correlation coefficients. However, previous research (e.g., Becker & Fahrbach, 1994; S. F. Cheung, 2001; Hafdahl, 2007) showed that this traditional GLS approach performed poorly comparing to the univariate approaches for the estimation of the pooled correlation matrix. Researchers have proposed different modifications of the original GLS method. For example, Becker and Fahrbach (1994) noted that using a mean correlation coefficient in Equation 2.3 was superior to the substitution using a sample correlation coefficient when the fixed-effects model was applied. Since the sample correlation coefficients contain measurement and sampling errors, the averaged correlation is a more reliable estimate of each population correlation (Furlow & Beretvas, 2005). Specifically, Becker and

11

Fahrbach (1994) used a sample sized weighted mean $\bar{r}_j$ , for $j = 1$ to $p*$, to substitute the population correlation coefficients in Equation 2.2 and Equation 2.3. The estimate of $\bar{r}_j$ was calculated as: $\bar{r}_j = \sum_{i=1}^{k} r_{ij} / n_i$, where $j = 1, 2, ..., p*$.

Define **r** as the $kp^*$ dimensional vector created by stacking up the correlation matrices from $k$ studies, for example, with $k = 5$ studies we stack the $6 \times 1$ **r**$_i$ s into the $30 \times 1$ outcome vector **r**:

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_1 \\ . \\ . \\ . \\ . \\ \mathbf{r}_5 \end{bmatrix}.$$  (2.4)

The $kp* \times kp*$ block-diagonal matrix **V** is formed by placing the $k$ **V**$_i$ matrices on the diagonal, that is, $\mathbf{\Sigma} = diag(\mathbf{V}_1, ..., \mathbf{V}_k)$, in this example

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{V}_5 \end{bmatrix}.$$  (2.5)

Next let **X** be the design matrix consisting of $k$ stacked $p* \times p*$ identity matrices. In this example

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ . \\ . \\ . \\ \mathbf{X}_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 \\ \hline ... & . & . & ..... \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$  (2.6)

To test the hypothesis of homogeneity of correlation matrices across $k$ studies, that is, to test $H_0 : \mathbf{\rho}_1 = \mathbf{\rho}_2 = ... = \mathbf{\rho}_k$ , use the test statistic

$$Q = \mathbf{r}'[\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]\mathbf{r}.$$  (2.7)

Under the null hypothesis, $Q$ has approximately a chi-square distribution with $(k$-$1)p^*$ degrees of freedom (Becker, 1992; Hedges & Olkin, 1985). If the test is not significant, which means that the population correlation matrices are homogeneous, a pooled correlation matrix can be obtained by using the GLS method, specifically

$$\mathbf{r.} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^{-1}\mathbf{V}^{-1}\mathbf{r}. \tag{2.8}$$

The approximate variance-covariance matrix of the estimate is

$$\mathbf{V(r.)} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{2.9}$$

It can be used to construct large-sample confidence intervals for the elements of the correlation matrix.

The GLS approach can be extended using Fisher's $z$ transformation. The asymptotic covariance matrix of the transformed $z$ values can be estimated by formulas provided in Steiger (1980). Becker and Fahrbach (1994) gave the pooled matrix of $z$ values and they suggested that the performance of GLS using Fisher z scores was better than the GLS method using raw correlation coefficients in estimating the pooled correlation matrix. They also noted that the asymptotic covariance matrix of Fisher $z$ scores cannot be easily transformed into the asymptotic covariance matrix of correlation matrix. Fouladi (2000) proposed an alternative approach to model Fisher $z$ scores with their asymptotic covariance matrix directly. However, the interpretation of the parameter estimates gets really complicated because the metric is in Fisher $z$ scores rather than correlations. To focus on the comparison of the methods on synthesizing raw correlation coefficients and modeling pooled correlation matrix, the Fisher-$z$ score transformation is not explored in this dissertation.

Sometimes we assume that there is variation among the population correlation vectors across studies. The between-studies model reflects this variation. For a fixed-effects model, which assumes that an underlying population correlation matrix is common to all studies, the model is $\boldsymbol{\rho}_i = \boldsymbol{\rho}$. For a random-effects model, which assumes that the population parameters are random samples from a hyper-population, the model is $\boldsymbol{\rho}_i = \boldsymbol{\mu}_\rho + \mathbf{u}_i$, where $\boldsymbol{\mu}_\rho$ denotes the mean population correlation, $\mathbf{u}_i$ reflects between-study variation, and $\mathbf{u}_i \sim N_{p^*}(\mathbf{0}, \mathbf{T})$, with $\mathbf{T}$ being the between-studies covariance

matrices. Thus, combining the within-study and between-studies models, we get the fixed-effects model

$$\mathbf{r}_i = \mathbf{\rho} + \mathbf{e}_i,$$

and the random-effects model

$$\mathbf{r}_i = \mathbf{\mu}_\rho + \mathbf{u}_i + \mathbf{e}_i.$$

One pervasive problem in the synthesis of correlation matrices is missing data. It is not likely for every study to report each element in the correlation matrix. Thus, dealing with missing data is an important concern in applying these multivariate methods. Becker and Schram (1994) suggested substituting sample estimates (such as means across studies) for missing population correlations in equation 2.3. Then one can modify equations 2.4-2.6 by essentially deleting the columns and rows of the matrices $\mathbf{r}$, $\mathbf{X}$ and $\mathbf{V}$ that correspond to missing values. Researchers have examined the missing data problem in synthesizing correlation matrices and given suggestions (e.g., Becker & Schram, 1994; Furlow & Beretvas, 2005). Missing not at random data seems to affect the precision of the correlations as well as the model parameters at the second stage of MASEM using either the univariate approaches or the multivariate approaches of MASEM. Since there is little evidence that missing data will particularly impact certain multivariate approach in the previous research, in this dissertation, the missing data issue is not considered for the methods comparison.

*Two-Stage Structural Equation Modeling (TSSEM) Method*

Cheung and Chan (2005) proposed a two-stage approach for meta-analytic structural equation modeling (TSSEM). At the first stage, correlation matrices are tested for homogeneity and the pooled correlation matrix and its asymptotic covariance matrix are estimated. At the second stage, the pooled correlation matrix is analyzed and its asymptotic covariance matrix is used as the weight matrix in a structural equation analysis with the asymptotic distribution-free (ADF) method as the estimation method.

To test the homogeneity of correlation matrices and estimate a pooled correlation matrix, the authors apply a multiple-group confirmatory factor analysis (CFA) method (see Jöreskog & Sörbom, 1993). The models below are written based on the observed variables, i.e., raw data. Correlation matrices are more often reported than covariance

14

matrices in meta-analysis. Hunter and Schmidt (1990) pointed out that comparing the standardize metric (correlation matrix) across studies is more interpretable and preferable in many situations. The correlation matrix of these observed variables is of interest in this project.

Let $\mathbf{x}_i$ be a $p_i \times n_i$ observed vector of the raw data and $\Sigma_i$ be its $p_i \times p_i$ population covariance matrix in the *i*th study, where $p_i$ is the number of observed variables, $i = 1, \ldots,$ *k* and $n_i$ is the sample size for study *i*. Since not every study reports *p* variables, $p_i$ can be less than or equal to *p*.

The covariance matrix $\Sigma_i$ can be decomposed into the matrices of standard deviations and correlations,

$$\Sigma_i = \mathbf{D}_i \mathbf{R}_i \mathbf{D}_i', \tag{2.10}$$

where $\mathbf{D}_i$ is the $p_i \times p_i$ diagonal matrix of standard deviations with $\mathbf{D}_i'$ its transpose, and $\mathbf{R}_i$ is the $p_i \times p_i$ correlation matrix in the *i*th study where $Diag[\mathbf{R}_i] = \mathbf{I}$.

In a confirmatory factor analysis (CFA), the observed vector $\mathbf{x}_i$ is hypothesized to be a function of $\mathbf{\Lambda}_i, \boldsymbol{\xi}_i$ and $\boldsymbol{\delta}_i$. Specifically, $\mathbf{x}_i = \mathbf{\Lambda}_i \boldsymbol{\xi}_i + \boldsymbol{\delta}_i$, where $\mathbf{\Lambda}_i$ is the factor loading matrix, $\boldsymbol{\xi}_i$ is the latent factor vector and $\boldsymbol{\delta}_i$ is the error vector. By assuming that the errors are uncorrelated and have a mean of zero, the covariance matrix $\Sigma_i$ of the observed variables $\mathbf{x}$ can be expressed as

$$\Sigma_i = \mathbf{\Lambda}_i \boldsymbol{\Phi}_i \mathbf{\Lambda}_i' + \boldsymbol{\Theta}_i, \tag{2.11}$$

where $\boldsymbol{\Phi}_i$ is the latent factor covariance matrix, and $\boldsymbol{\Theta}_i$ is the error covariance matrix. By imposing constraints on model parameters, Equation 2.11 can be used to test hypotheses about the covariance matrix of the observed variables through the latent factor covariance matrix (Raykov, 2001). For example, suppose $\mathbf{\Lambda}_i = \mathbf{I}_{p_i}$ and $\boldsymbol{\Theta}_i = \mathbf{0}_{p \times p}$, where $\mathbf{I}_p$ is the $p \times p$ identity matrix and $\mathbf{0}_{p \times p}$ is the $p \times p$ zero matrix, then Equation 2.11 simplifies to a special case: $\Sigma_i = \boldsymbol{\Phi}_i$. $\tag{2.12}$

Raykov (2001) pointed out the model 2.12 "can be seen as resulting after $q = p$ dummy latent variables $\xi_1, \xi_2, \ldots, \xi_p$ are introduced, which equal each a corresponding

observed variable, namely, $\xi_i = x_i$ ($i = 1,\ldots, p$)." (p. 226). It is a basic model in testing multivariable covariance structure hypotheses because the observed covariance matrix $\Sigma$ is equal to the latent covariance matrix $\Phi$ after these constraints and each of its elements in the latent construct is now accessible to the researcher. Figure 2.1 depicts this 'null' model.

With a further extension of the specifications, Cheung and Chan (2004) specify that if $\Lambda_i$ is a $p_i \times p_i$ diagonal matrix, $\Phi_i$ is a $p_i \times p_i$ standardized matrix, and $\Theta_i$ is a $p_i \times p_i$ zero matrix, then Equation 2.10 is equivalent to Equation 2.11.

Under the above constraints the factor correlation matrix is equivalent to the observed correlation matrix, and testing hypotheses about the factor correlation matrix is equivalent to testing the hypotheses about the correlation matrix for the observed variables. We can see that the amount of information in the model across $k$ studies is $\sum_{i=1}^{k} \dfrac{p_i(p_i + 1)}{2}$, where $p_i$ is the number of variables in the $i$th study. The number of free parameters in estimating the correlation across different studies is $\sum_{i=1}^{k} \dfrac{p_i(p_i - 1)}{2}$, so the number of free parameters in estimating the standard deviation matrices is $\sum_{i=1}^{k} p_i$ . Thus, the degree of freedom equals the amount of information minus the number of free parameters, i.e, $df = \sum_{i=1}^{k} \dfrac{p_i(p_i + 1)}{2} - [\sum_{i=1}^{k} \dfrac{p_i(p_i - 1)}{2} + \sum_{i=1}^{k} p_i] = 0$. Thus fitting the model in equation 2.10 to the observed data will lead to a perfect fit with $df = 0$, and the parameter estimates will equal to the observed data. Here the correlation matrix is used as a covariance matrix. Since the standard deviation matrix $\mathbf{D}_i$ contains information about scale, the estimate of the $\mathbf{R}_i$ matrix will presumably be the same using either the correlation or covariance matrix as input matrix (Cheung & Chan, 2005, p. 45).
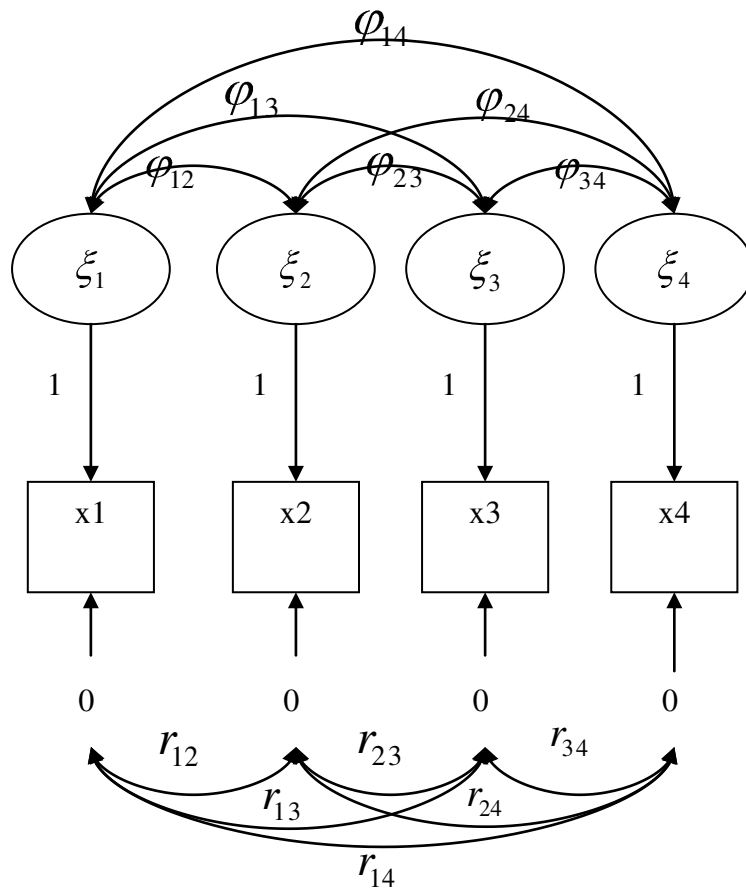
Figure 2.1 The 'null' model where the observed correlation matrix is equal to the latent correlation matrix

Assuming that all population correlation matrices are from the same population, that is, $H_0 : \boldsymbol{\rho}_1 = \boldsymbol{\rho}_2 = ... = \boldsymbol{\rho}_k$, multiple-group SEM techniques can be used to obtain the pooled correlation matrix. Multiple-group SEM (also called multiple-sample SEM) is a technique to address the question that whether values of model parameters vary across groups (Kline, 2005). Usually in MASEM an SEM computer program is used to perform this analysis which simultaneously estimates a model across all samples. Through the specification of cross-group equality constraints, group differences on any individual parameter or set of parameters can be tested. First, constrain the factor correlation matrices $\boldsymbol{\Phi}_i$ to be equal across studies. The number of between-study constraints is then

$$\sum_{i=1}^{k} \frac{p_i(p_i - 1)}{2} - \frac{p(p-1)}{2}.$$

The test statistic is asymptotically distributed as a chi-square with degrees of freedom

$\sum_{i=1}^{k} \frac{p_i(p_i - 1)}{2} - \frac{p(p-1)}{2}$. When all studies have all p variables, this equation turns into:

$(k-1)p(p-1)/2 = (k-1)p^*$. Note that the degrees of freedom is the same as that of the multivariate $Q$ test as described in Equation 2.7. A chi-square difference test is conducted to compare this model with constraints to the model without constraints (the model with $df = 0$). If the fit of the constrained model is not significantly worse than the model without constraints, the correlation matrices will be treated as homogeneous. Then a pooled correlation matrix estimate $\mathbf{R}$ (i.e., the reproduced correlation matrix) and its asymptotic covariance matrix $\mathbf{V_R}$ can be computed.

   Next, the pooled correlation matrix $\mathbf{R}$ and its asymptotic matrix $\mathbf{V_R}$ are used as input in the second step to do path analysis, confirmatory factor analysis or to compute structural equation models. The pooled correlation matrix $\mathbf{R}$ is a $p \times p$ matrix and its asymptotic matrix, $\mathbf{V_R}$, is a $p(p-1)/2 \times p(p-1)/2$ or $p^* \times p^*$ matrix. For example, if $\mathbf{R}$ is a $4 \times 4$ matrix, then the $\mathbf{r}$ vector has $p(p-1)/2 = 6$ distinct correlation coefficients and is created by stringing out the lower triangle of the $\mathbf{R}$ matrix. Thus, the asymptotic matrix $\mathbf{V_R}$ incorporates the variances and covariances among these 6 distinct correlations, and results in a matrix of dimension $6 \times 6$. The total sample size $N$ equals $\sum_{i=1}^{k} n_i$. By using an ADF estimation method, a proposed model on the pooled correlation matrix can be tested. The model is $\mathbf{R} = \mathbf{R}(\mathbf{\Omega})$, and $\mathbf{R}(\mathbf{\Omega})$ represents a correlation structure model with parameters $\mathbf{\Omega}$ (for example, $\mathbf{R}(\mathbf{\Omega})$ can be a correlation structure of a path analysis model or a full structural model). The fit function is then

$$F(\mathbf{\Omega}) = (\mathbf{r} - \mathbf{r}(\mathbf{\Omega}))^T \mathbf{V_R}^{-1} (\mathbf{r} - \mathbf{r}(\mathbf{\Omega})), \tag{2.13}$$

where $\mathbf{r}$ and $\mathbf{r}(\mathbf{\Omega})$ are the $p^* \times 1$ vectors of $p^* = p(p-1)/2$ elements obtained by stringing out the lower triangular elements of the correlation matrices $\mathbf{R}$ and $\mathbf{R}(\mathbf{\Omega})$, respectively. The weight matrix $\mathbf{V_R}$ is the asymptotic covariance matrix of the pooled

correlation matrix **R**. Equation 2.13 is equivalent to the overall $Q$ test in GLS method

when $\mathbf{R}(\mathbf{\Omega}) = \mathbf{r.}$, while the weight matrices are different for these two methods.

Cheung (2002) gave three reasons for using ADF, instead of maximum likelihood (ML)

or GLS as the estimation method at the second stage of the TSSEM approach. First, since

the total sample size $N = \sum_{i=1}^{k} n_i$ is used in the second stage and it is generally large

enough for ADF to lead to correct inferences. One reason for using the total sample size

in the second stage is that it will increase the statistical power (Hedges & Pigott, 2001).

Second, the sampling variation among studies is reflected in the weight matrix $\mathbf{V_R}^{-1}$, with

less (more) weight being given to those with large (small) amounts of sampling variation.

Third, since the statistical theory of structural equation modeling is based on the

distribution of covariance instead of correlation, it is generally not appropriate to apply

covariance distribution theory without adjustment (Cheung, 2002, p. 53). Thus, ADF is

more appropriate as the estimation method when a correlation matrix instead of a

covariance matrix is used as input for stage 2 analyses. The minimum of $(N-1)F(\mathbf{\Omega})$ is

asymptotically distributed as chi-square with ($p*-q$) degrees of freedom, where $N$ is the

total sample size and $q$ is the number of free parameters.

*GLS Method and TSSEM Method Comparisons*

   Both GLS and TSSEM methods are multivariate methods which model the

dependence among correlation coefficients and both methods involve correlations among

the correlation coefficients and the asymptotic variance matrix of the pooled correlation

matrix. Cheung (2002) pointed out that the main differences between these two methods

are due to their statistical methods. For example, with the GLS method, Olkin and

Siotani's (1976) equation given in equation 2.3 above is used as the asymptotic

covariance matrix of the correlation matrix based on the large sample distribution theory

for correlations, and a GLS approach is used in the second stage to estimate the pooled

correlation matrix. The TSSEM method is based on covariance distribution theory to

estimate the pooled correlation matrix by treating the diagonals as parameters that need to

be estimated. As pointed out by Cudeck (1989), analyzing a correlation matrix as a

covariance matrix may produce incorrect chi-square statistics, goodness-of-fit indices and

incorrect standard errors. One approach to analyzing the correlation matrix directly with a correlation structure is to use ADF estimation based on Equation 2.13 with which a modification of Browne (1984)'s ADF method is applied.

## Comparison of Methods for Meta-analysis of Correlation Matrices

To date, several research papers have examined the performance of the different univariate and multivariate methods of synthesizing correlation matrices (e.g., Becker & Fahrbach, 1994; S. F. Cheung, 2000; Cheung & Chan, 2005; Hafdahl, 2007; Hafdahl, 2008). Cheung and Chan (2005) found that the univariate methods (Hedge & Olkin, 1985; Hunter & Schmidt, 1990, 2004) control Type I error well in testing the homogeneity of the correlation matrices and the biases of parameter estimates are small, but they are too liberal in controlling the Type I error in the modeling process. Among the multivariate methods, different methods of estimating the pooled correlation matrix result in different values of the correlation matrix as well as different asymptotic covariance matrices. Becker and Fahrbach (1994) noted that the GLS method performed poorly in small to medium samples (say, sample sizes smaller than 250). Since the sample correlation contains both measurement and sampling errors, substituting the sample correlations for the population correlations in the variance-covariance matrix may lead to the inefficiency of the estimates of covariances between correlation coefficients. Becker and Fahrbach (1994) used the simple average of the relevant correlations for estimates of the population correlations in the covariance matrix and found that the precision of estimation improved. Similarly, S. F. Cheung (2000) used a sample size weighted correlation for the population correlation in formula 2.3. Both pieces of research found that the modified estimate of the population correlations outperform the traditional GLS approach. Specifically, S. F. Cheung (2000) found that the modified GLS approach has higher power in the chi-square test to reject the assumption of homogeneity of the correlation matrices when they were actually heterogeneous. Furlow and Beretvas (2005) did a simulation study to compare the performance of different methods for synthesizing correlation matrices for MASEM under different patterns of missing data. They compared the estimates of correlations and SEM parameters and model fit. They

used a variance weighted correlation as a substitute for the population correlation in the variance-covariance matrix and found that the modified method performed better for estimation of parameters and gave more accurate model rejection rates.

Hafdahl (2007) examined several methods for synthesizing correlation matrices, which included univariate methods and multivariate methods (specifically, GLS methods). He substituted sample size weighted correlations for the population correlations in the $\mathbf{V}_i$ matrix. Also he expanded the conditions Becker and Fahrbach (1994) examined with some extensions. For example, a correlation matrix of $p = 4$ variables instead of $p = 3$ was used and the within-study sample sizes were varied. The sample sizes were drawn from a positively skewed distribution which was considered typical in research syntheses (Osburn & Callender, 1992). He concluded that sampling error in observed correlations is the most important reason for the GLS method to perform poorly, as reported in Becker and Fahrbach (1994). He wrote "Fisher's *z*-transformation improved performance and was necessary to maintain multivariate homogeneity rejection rates at nominal --- due perhaps to homogeneity tests' greater reliance on normality via their quadratic forms --- but was less effective than the estimated population correlation refinement." (Hafdahl, 2007, p. 198)

Among the aforementioned research, two studies of GLS methods (Becker & Fahrbach, 1994; Hafdahl, 2007) compared methods of synthesizing correlation matrices under the fixed-effects model with homogeneous data. Two studies (Cheung & Chan, 2005; S. F. Cheung, 2000) examined the homogeneity test with both homogeneous and heterogeneous data. One study (Hafdahl, 2008) examined the meta-analytic methods for correlation matrices using heterogeneous data. They all assumed a fixed-effects model for their analyses. That is, a single population correlation matrix is assumed to underlie all the correlation matrices. However, this assumption will be challenged in practice given that studies are often heterogeneous. Thus random-effects models may be more realistic in practice.

A random-effects model for the synthesis of correlation matrices has been described by Becker (1992, 1995). Recently, Prevost et al. (2007) considered different methods for allowing for correlations between correlation coefficients, including generalized least squares (GLS), maximum marginal likelihood, and Bayesian

approaches using a random-effects model. They examined the quantities of the mean population correlation matrix, contrasts between the mean correlations and the prediction of a correlation matrix in a new study. They concluded that compared to the univariate methods, considering the correlations among the correlation coefficients would have little impact on the individual estimates of the correlation coefficients. However, for the composite measure of the correlation matrix or overall hypothesis testing about the correlation matrix and the successive analyses of the correlation matrix, multivariate methods can have notable impact. In particular, the authors compared the correlation coefficients estimates from eight estimation methods with a real example. They examined random-effects method using non-iterative GLS, a random-effects method using maximum marginal likelihood, and a random-effects method using Markov chain Monte Carlo (MCMC). They concluded that the GLS methods (both iterative and noniterative), "can lead to inappropriate answers that are due in part to the lack of a check on the appropriateness of the estimate of the between-study random-effects covariance." They recommended a model-based maximum marginal likelihood approach or a full Bayesian analysis.

Moderator analysis has not been examined in most of the studies discussed above, though Hafdahl (2007) pointed out that several complex models would be readily appropriate, for example, fixed-effects models with study characteristics, random-effects models and mixed-effect models. These models were examined using treatment-type effect sizes in Kalaian and Raudenbush (1996).

## Synthesizing Covariance Matrices in Meta-Analytic SEM

Beretvas and Furlow (2006) pointed out that it is generally not appropriate to use a pooled correlation matrix as the input for the second stage in MASEM, since common SEM programs are all based on the analysis of a covariance matrix. Cudeck (1989) delineated the potential problems of applying statistical theory applicable for covariance structure analysis to correlation matrices. Specifically, he pointed out that the chi-square test of model fit (and thus associated fit indices) and the standard errors associated with

path coefficients would be impacted when the model under analysis is not scale-invariant. Beretvas and Furlow (2006) explained the concept of 'scale invariant' as follows:

"The concept of scale invariance can be best explained through the use of a confirmatory factor analysis (CFA) example, although the concept applies to the family of models fitting within SEM. First, the covariance matrix of observed variables being factor-analyzed can be decomposed into the model's parameters. Thus, under the general factor analysis model, the covariance matrix, $\Sigma(\gamma)$, based on a vector of parameters, $\gamma$, can be decomposed into a function of the factor loadings' matrix, $\Lambda$, the factor covariance matrix, $\Phi$, and the matrix of errors, $\Psi$, such that $\Sigma(\gamma) = \Lambda\Phi\Lambda' + \Psi$ . The definition of scale invariance is that "any rescaling of a covariance matrix, result in another covariance matrix that also satisfies the model" (Cudeck, 1989, p. 319). Specifically, a scale-invariant will result if there exists a diagonal matrix, D (with no zeros along its diagonal), such that a different covariance matrix, , for the vector of rescaled parameters $\gamma^*$, exists where $\Sigma(\gamma^*) = D\Sigma(\gamma)D$ " (p. 155).

If, however, a model is scale invariant then only the standard error estimates for the parameters that are not 'scale free' would be affected. Here, 'scale free' refers to the property that the values of the parameter estimates remain the same regardless of the rescaling of the covariance matrix (Beretvas & Furlow, 2006, p. 155). Given the fact that many of the models examined by MASEM researchers are not scale invariant, Beretvas and Furlow (2006) argued that the model-fit criteria based on the chi-square statistic as well as the standard errors of parameters which are not scale free might not be correct. With this concern in mind, they applied the synthesis of standard deviations of the observed variables along with the synthesis of correlation matrices. Therefore, an approximation to a synthesized covariance matrix was obtained from synthesized correlations and synthesized standard deviations.

It is not known how practical this method will be in applied research since another layer of information is needed for implementation of this approach, that is, the standard deviations of the observed variables. Beretvas and Furlow (2006) also discuss the potential problems using the pooled correlation matrix in standard SEM programs (as

pointed out in Cudeck, 1989). Even with the ability of some programs (e.g., LISREL, Mx) to correct for the standard errors of the parameters, the model-fit criteria based on chi-square test may still not correct for the non-scale-invariant models. Cheung and Chan (2009) also proposed a two-stage approach to synthesizing covariance matrices, which is the extension of the Cheung and Chan (2005) TSSEM method.

Though different methods of synthesizing correlation matrices have been proposed by many researchers, performance differences among these methods for further analyses of various functions of the synthesized correlation matrix are not well studied. Becker (1992) proposed a procedure for analyzing the pooled correlation matrix to obtain linear models. Cheung and Chan (2005) proposed techniques for analyzing confirmatory factor analyses and structural equation models. How these methods of meta-analyzing correlation matrices differ at the stage of the modeling process is my primary interest in this research.

Specifically, path analyses use the correlation matrix as the input. Becker (1992) described a direct computation method which produces estimates of path coefficients and standard errors based on the average correlation matrix. This approach may not have the potential problems associated with applying correlation matrices instead of covariance matrices with SEM software packages. The comparison of methods for this modeling process versus the MASEM approach is described in chapter III.

# CHAPTER III

# METHODOLOGY

In the previous chapter, methods of pooling correlation matrices across independent studies were described. Recently developed methods for meta-analytic structural equation modeling were also introduced. My primary goals in this dissertation are to apply and compare the multivariate methods of synthesizing correlation matrices and meta-analytic path analyses using a direct computation method (Becker, 1992) and the TSSEM method of Cheung and Chan (2005). I will examine the overall homogeneity tests of these methods at the first step and compare the parameter estimates and standard errors of the path coefficients from these two methods at the second step. Suggestions regarding the limitations and advantages of each method will be given. The method of estimating path coefficients based on an average correlation matrix from the random-effects model will be explored as well, given the fact that a random-effects model is probably more appropriate in many situations in practice. Since the SEM method with maximum likelihood estimation as the second-step estimation method is used in many meta-analytic path analyses, it will also be discussed to show the comparison of the parameter estimates with those from the GLS direct computation methods and the original TSSEM methods.

## MASEM using GLS Approach and TSSEM Approach

Path analysis is an important analytical tool in testing and estimating causal relationships among observed variables. With the development of the general structural equation modeling techniques, path analysis has been widely applied and studied in the fields of education, psychology, politics and economics etc (Bollen, 1989). Recently, path analysis has been used in many research syntheses (e.g. Brown, Tramayne, Hoxha, Telander, Fan & Lent, 2008; Fried, Shriom, Gilboa & Cooper, 2008; Zimmerman & Darnold, 2009). Meta-analytic path analysis has been used by researchers to test their

hypotheses in the theoretical structures and whether the hypothesized causal models are consistent with the empirical findings from different studies. Shadish (1996) noted the importance of using causal mediating models like path analysis models to address research questions about theory development in meta-analysis. Though other types of meta-analytical structural equation modeling techniques has been employed in substantive research, for example, meta-analytic confirmatory factor analysis (e.g., Cheung & Au, 2006), so far meta-analytic path modeling is still the most frequently used techniques by the applied researchers. In this dissertation, meta-analytic path modeling is used interchangeably with the term MASEM.

Path analysis can be conducted as linear regression model (e.g., Pedhazur, 1997). More popularly, path analysis is treated as a subset of techniques in the general framework of structural equation modeling. Below I talk about estimating path models using both techniques.

### Direct Computation Method

Pedhazur (1982, 1997) described a way of doing path analysis using regression models. Schmidt, Hunter, and Outerbridge (1986) applied this method to do a path analysis using the pooled correlation matrix. Path modeling is based on the correlations among the observed variables and uses the correlation matrix instead of the covariance matrix as an input matrix. S. F. Cheung (2000) pointed out that it is specifically developed for analyzing correlations.

The idea of using the regression method for obtaining path analysis results is to estimate a set of regression models for the individual paths in the path model. Consider model 1 in Figure 3.1 as an example.
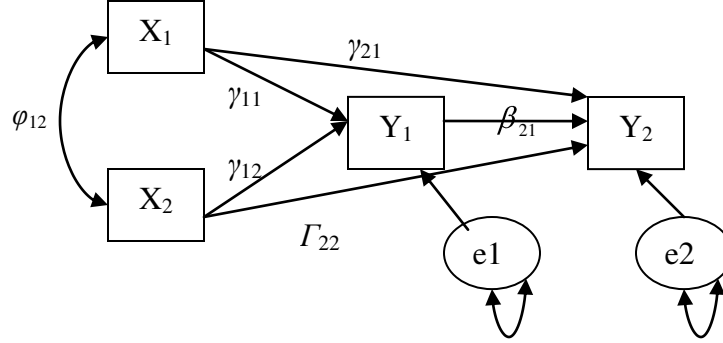
Figure 3.1 Example path analysis model 1

The model specified above is a recursive and just-identified (saturated) path model which has degrees of freedom equal to zero. A recursive model has two basic features: one is that all the causal effects are unidirectional and have no feedback loops between each other; the other is that the disturbances (error variances) are not correlated (Kline, 2005, p. 104). Having a just-identified model means that the amount of observed information in the model is equal to the number of parameters to be estimated. To get a unique solution for each parameter, the model has to be at least just-identified ($df \geq 0$). Recursive path models are always identified (see Bollen, 1989, pp. 95-98).

Let $s$ be the number of endogenous variables and $t$ be the number of exogenous variables. (The 'endogenous variables' refer to the variables to be viewed as outcome, like 'dependent variables', while 'exogenous variables' refer to variables acting as predictors, like 'independent variables' in linear models). The system of structural equations representing the model in figure 3.1 can be written as

$$\mathbf{y} = \mathbf{By} + \mathbf{\Gamma x} + \zeta \tag{3.1}$$

where $\mathbf{y}$ is a $s{\times}1$ vector of observed endogenous variables, $\mathbf{x}$ is a $t{\times}1$ vector of exogenous variables, $\mathbf{B}$ is a $s{\times}s$ coefficient matrix that relates endogenous variables to each other, $\mathbf{\Gamma}$ is a $s{\times}t$ matrix of coefficients that relates endogenous variables to exogenous variables, and $\zeta$ is a $s{\times}1$ vector of disturbances, where $\text{cov}(\zeta) = \mathbf{\Psi}$ is the $s{\times}s$ covariance matrix of the disturbances. The $t{\times}t$ covariance matrix for the exogenous variable is $\text{cov}(\mathbf{x}) = \mathbf{\Phi.}$

For example, the number of pieces of observed information in model 3.1 is 4(4+1)/2 = 10, and the parameter set being estimated in model 3.1 is $\{\phi_{11}, \phi_{22}, \phi_{12}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, \beta_{21}, \zeta_1, \zeta_2\}$. This model is a simple model with both mediating effects (e.g., the effect of $X_2$ on $Y_2$ via $Y_1$) and direct effects (e.g., the effect of $Y_1$ on $Y_2$). Also, it is a model that will perfectly fit the data. It is easy to change it from a just-identified model to an over-identified model (e.g., by removing the path from $X_2$ to $Y_2$ or any other path or paths, see Figure 3.2).
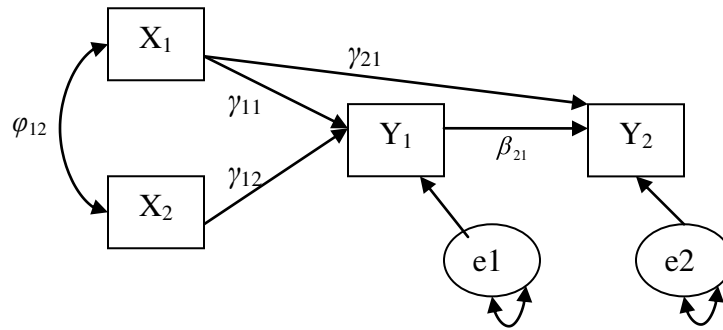


Figure 3.2 Example path analysis model 2

Of course, in real practice the formation of a path model needs to be built on a reasonable theoretical basis. The regression method will evaluate a set of linear models to establish this path analysis. For the example showed in Figure 3.1, we see that the following linear equations are examined:

$$Y_1 = \gamma_{11}X_1 + \gamma_{12}X_2 + e_1$$

and $\quad Y_2 = \gamma_{21}X_1 + \gamma_{22}X_2 + \beta_{21}Y_1 + e_2,$

For the over-identified model shown in Figure 3.2, the linear equations are

$$Y_1 = \gamma_{11}X_1 + \gamma_{12}X_2 + e_1,$$

and $\quad Y_2 = \gamma_{21}X_1 + \beta_{21}Y_1 + e_2.$

Though Pedhazur (1982, 1997) showed how to formulate path analysis models as linear regression models, confidence intervals for the parameter estimates are not easily computed. Becker (1992) proposed a method that gives a large sample estimate of the variance-covariance matrix of the parameter estimates. The method uses the multivariate delta method to obtain the asymptotic distribution of the standardized regression

coefficients. Thus, the confidence intervals and the statistical significance of the standardized regression coefficients can be obtained. Below is a brief summary of this method. (See these derivations in Becker, 1992, p. 358.)

Let $X_0$, $X_1$,…, $X_p$ be random variables with a multivariate normal distribution and population correlation matrix **P**. Partition **P** as

$$\mathbf{P} = \begin{bmatrix} 1 & \mathbf{P}_{10} \\ \mathbf{P}_{01} & \mathbf{P}_{11} \end{bmatrix},$$

so that the correlations of the predictors with the outcome or dependent variable are contained in $\mathbf{P}_{01}$, and the correlations among the predictors are in $\mathbf{P}_{11}$. Then the population standardized regression coefficients $\boldsymbol{\beta} = (\beta_1,...,\beta_p)'$ of the regression of $X_0$ on $X_1$,…, $X_p$ can be expressed as $\boldsymbol{\beta} = \mathbf{P}_{11}^{-1}\mathbf{P}_{01}$. Suppose **R** is the sample estimate of **P** and is partitioned as

$$\mathbf{R} = \begin{bmatrix} 1 & \mathbf{R}_{10} \\ \mathbf{R}_{01} & \mathbf{R}_{11} \end{bmatrix},$$

then $\mathbf{b} = (b_1,...,b_p)$ is the sample estimate of $\boldsymbol{\beta}$ and is given by

$$\mathbf{b} = \mathbf{R}_{11}^{-1}\mathbf{R}_{01}.$$

The above expression is similar to the formula for standardized regression coefficients such as that given in Cooley and Lohnes (1971).

Let $\mathbf{r.} = (r_{\cdot 01}, r_{\cdot 02},...,r_{\cdot 0p}, r_{\cdot 12},...,r_{\cdot(p-1)p})'$ and

$\boldsymbol{\rho} = (\rho_{01}, \rho_{02},...,\rho_{0p}, \rho_{12},...,\rho_{(p-1)p})'$ denote the vector of $p(p+1)/2$ nonredundant sample estimates and population correlations obtained by stringing out the lower triangle of the **R** and **P** matrices, respectively.

Suppose **R** is a sample estimate of the correlation matrix **P** with an asymptotic distribution. As the sample size $N \rightarrow \infty$, the distribution is given by $\sqrt{N}(\mathbf{r.} - \boldsymbol{\rho}) \sim N(\mathbf{0}, \mathbf{H}(\boldsymbol{\rho}))$, where $\mathbf{H}(\boldsymbol{\rho})$ is a function of $\boldsymbol{\rho}$, the asymptotic normal distribution of the standardized regression coefficient estimate **b** is then given by

$$\sqrt{N}(\mathbf{b} - \boldsymbol{\beta}) \sim N(\mathbf{0}, \mathbf{M}(\boldsymbol{\rho})),$$

where the function $\mathbf{M}(\boldsymbol{\rho})$ is the product of matrix **A** and $\mathbf{H}(\boldsymbol{\rho})$, $\mathbf{A} = \mathbf{A}(\boldsymbol{\rho}) = (a_{ia})$,

$$a_{ia} = \begin{cases} \rho^{ij} & \text{if } \rho_\alpha = \rho_{0j} \in \mathbf{P}_{01} \\ -\sum_{j=1}^{p} (\rho^{ki}\rho^{kj} + \rho^{lj}\rho^{li})\rho_{j0} & \text{if } \rho_\alpha = \rho_{kl} \in \mathbf{P}_{11} \end{cases}$$

and $\rho^{ij}$ are the elements of $\mathbf{P}^{-1}$ (see equation 19 in Becker, 1992, p. 358). Then the large sample multivariate normal approximation to the distribution of $\mathbf{b}$ is

$\mathbf{b} \sim \mathbf{N}(\boldsymbol{\beta}, \mathbf{S(r.)})$,

where $\mathbf{S(r.)} = \mathbf{A(r.)H(r.)A(r.)'}/N = \mathbf{A(r.)V(r.)A(r.)'}$; $\mathbf{A(r.)}$ and $\mathbf{H(r.)}$ are defined by replacing each population correlation in $\mathbf{A}$ and $\mathbf{H}$ by the corresponding sample correlation, and $\mathbf{V(r.)} = \mathbf{H(r.)}/N$ is the large sample covariance matrix of $\mathbf{r.}$. Also the confidence intervals for the standardized regression coefficients can be obtained. A $100(1-\alpha)$ percent confidence interval for $\beta_i$ is given by

$$b_i - z_{\alpha/2}\sqrt{s_{ii}} \le \beta_i \le b_i + z_{\alpha/2}\sqrt{s_{ii}},$$

where $s_{ii}$ is the $i$th diagonal element of $\mathbf{S(r.)}$ and $z_\alpha$ is the $100\alpha$ percent critical value of the standard normal distribution. The corresponding test of statistical significance of $\beta_i$ uses the statistic

$$T_i = b_i / \sqrt{s_{ii}} .$$

It is possible to test if a specified path model is consistent with the pattern of the inter-correlations among the variables. Below I will introduce briefly the overall model test for a path analysis model using regression analysis.

Generally, tests of over-identified models are performed by using properties of the observed and the reproduced correlation matrices among the variables under consideration. (In meta-analytic path modeling, the pooled correlation matrix from the meta-analysis is treated as the observed matrix).Essentially, determinants of these matrices are used to calculate a chi-square statistic with degrees of freedom equal to the number of over-identifying constraints.

In order to test an over-identified model, the first step is to calculate $R_G^2$, which is defined as a generalized squared multiple correlation. Specht (1975) showed that for a fully recursive model: $R_G^2 = 1 - (1 - R_1^2)(1 - R_2^2)...(1 - R_c^2)$, where $R_i^2$ is the ordinary square multiple correlation coefficient of the $i$th equation in a fully recursive system, $i = 1$,

2,…,$c$. For an over-identified model (e.g., the model in Figure 3.2), one can calculate a statistic similar to $R_G^2$:

$$R_M^2 = 1 - (1 - R_{1r}^2)(1 - R_{2r}^2)...(1 - R_{cr}^2)$$

Here $R_M^2$ is calculated in the same manner as $R_G^2$, except that the $R_{ir}^2 s$ are based on the restricted or constrained models in which some of the paths are missing comparing to the fully recursive model, here $ir = 1r,…cr$, for the number of equations in the restricted model ('$r$' means 'restricted' here in the subnotations). Therefore, $R_M^2$ can take values between zero and $R_G^2$. A measure of goodness of fit for an over-identified model is

$\frac{1 - R_G^2}{1 - R_M^2}$, and for large samples, this measure can be tested for significance as

$$W = -(N - d)\ln(\frac{1 - R_G^2}{1 - R_M^2}),$$

where $N$ is the sample size and $d$ is the number of constraints. The statistic $W$ has an approximate $\chi^2$ distribution with $df = d$ (Pedhazur, 1982, p. 619). The model is rejected if the $\chi^2$ test is significant at the selected significance level, suggesting that the model does not fit the data adequately.

When different models are of concern, and the models are nested, tests can also be employed to see which model fits the data better (Specht, 1975, p. 125). The two models are fitted to the data and a chi-square test statistic is computed based on the two models' generalized multiple correlations. If the test is significant, the model with the larger number of parameters will be preferred to the one with the smaller number of parameters at the selected significance level.

### Structural Equation Modeling Method

Path analysis is generally treated as a special case of SEM in which only a structural model is considered. That is, a single indicator is used for each variable studied and the causal relationships among them are modeled. Structural equation modeling seeks to describe the means, variances, and covariances of a set of variables in terms of a

smaller number of "structural parameters" (Kaplan, 2000, p. 13). Equation 3.1 described

the structural form of the structural equation model. It can be rewritten as

$$\mathbf{y} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{\Gamma}\mathbf{x} + (\mathbf{I} - \mathbf{B})^{-1}\mathbf{\zeta} \, ,$$

where $\mathbf{I}$ is an identity matrix here, and $\mathbf{B}$, $\mathbf{\Gamma}$, $\mathbf{\zeta}$ are the same as described in equation 3.1.

Let $E(\mathbf{x}) = \mathbf{\mu_x}$ be the vector of means for $\mathbf{x}$, $cov(\mathbf{x}) = E(\mathbf{x'x}) = \mathbf{\Phi}$, and $E(\mathbf{\zeta}) = 0$. Then,

$$E(\mathbf{y}) = (\mathbf{I} - \mathbf{B})^{-1}E(\mathbf{x}) = (\mathbf{I} - \mathbf{B})^{-1}\,\mathbf{\mu_x},$$

$$E(\mathbf{y}, \mathbf{x}) = \mathbf{\Sigma} = \begin{bmatrix} E(\mathbf{yy'}) & E(\mathbf{yx'}) \\ E(\mathbf{x'y}) & E(\mathbf{x'x}) \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{I}-\mathbf{B})^{-1}(\mathbf{\Gamma}\mathbf{\Phi}\mathbf{\Gamma'}+\mathbf{\psi})(\mathbf{I}-\mathbf{B})^{'-1} & (\mathbf{I}-\mathbf{B})^{-1}\mathbf{\Gamma}\mathbf{\Phi} \\ \mathbf{\Phi}\mathbf{\Gamma'}(\mathbf{I}-\mathbf{B})^{-1} & \mathbf{\Phi} \end{bmatrix} \qquad (3.2)$$

Equation 3.2 tells us that the mean vector and covariance matrix can be represented by

the parameters of the model.

To estimate the model parameters, several estimation methods can be applied.

Among them, the most often used is the Maximum Likelihood (ML) method. It is the

default estimation method of some major SEM program (like LISREL, by Jöreskög and

Sorbom, 2000; and Mplus, by Muthén and Muthén, 1998, etc.). Generalized Least

Squares (GLS) is another popularly used estimation method. Both of these methods

assume that the data are a realization of a multivariate normal distribution and both are

scale free and scale invariant (Kline, 2005).

Let $X_s$ be a $p \times 1$ random vector and $\mathbf{\mu}_s$ be its mean vector and $\mathbf{\Sigma}_s$ be its

$p \times p$ population covariance matrix. The population covariance matrix $\mathbf{\Sigma}_s$ is the function

of the unknown model parameters $\mathbf{\Omega}$, which has dimension $q \times 1$, where $p$ is the total

number of observed variables (in path analysis, the $p$ variables include both endogenous

and exogenous variables), and $q$ is the number of free parameters. We assume

that $\mathbf{\Sigma}_s = \mathbf{\Sigma}_s(\mathbf{\Omega})$, where $\mathbf{\Sigma}_s(\mathbf{\Omega})$ can be a path analysis model, a CFA model or a full

SEM model (e.g., Bollen, 1989). To obtain an estimate of the parameter $\mathbf{\Omega}$, which is

denoted as $\hat{\mathbf{\Omega}}$ here, we generally minimize a discrepancy function $F(\mathbf{S}, \hat{\mathbf{\Sigma}}_s)$, where $\mathbf{S}$ is

the sample covariance matrix, and $\hat{\mathbf{\Sigma}}_s = \mathbf{\Sigma}_s(\hat{\mathbf{\Omega}})$ is the implied covariance matrix which is

based on the estimates of the model. This discrepancy function is a scalar and it measures

the distance between the sample covariance matrix $\mathbf{S}$ and the implied covariance $\hat{\Sigma}_s$.

Generally, the following function is used to obtain the estimated parameters:

$$F(\mathbf{S}, \hat{\Sigma}_s) = (\mathbf{s} - \boldsymbol{\sigma}(\hat{\boldsymbol{\Omega}}))' \mathbf{W}^{-1} (\mathbf{s} - \boldsymbol{\sigma}(\hat{\boldsymbol{\Omega}})), \tag{3.3}$$

where $\mathbf{s}$ and $\boldsymbol{\sigma}(\boldsymbol{\Omega})$ are vectors of dimension $p^* \times 1$ and $p^* = p(p+1)/2$. They are obtained by stringing out the elements in the lower triangles of the sample covariance $\mathbf{S}$ and implied covariance matrix $\hat{\Sigma}_s$. $\mathbf{W}^{-1}$ is a weight matrix of dimensions $p^* \times p^*$, and it is typically considered to be fixed, possibly estimated from the sample and a positive definite matrix (Browne, 1984).The minimum of the discrepancy function follows a chi-square distribution in large samples with appropriate selection of $\mathbf{W}^{-1}$(Cheung, 2002, p. 37). Cheung and Chan (2005) defined the weight matrix as the asymptotic covariance matrix of the pooled correlation coefficients estimated from their stage 1 analysis (see for the specific discrepancy function used). They proposed to use the ADF estimation method for parameter estimation in the second stage.

With different choice of the weight matrix, several discrepancy functions are defined. One is the general least squares (GLS) estimation method, with which one chooses $\mathbf{W}^{-1} = \mathbf{S}^{-1}$ .Thus the fit function in Equation 3.3 can be simplified as

$$F_{GLS} = \frac{1}{2} tr[\mathbf{S}^{-1}(\mathbf{S} - \boldsymbol{\Sigma}_s(\hat{\boldsymbol{\Omega}}))]^2 = \frac{1}{2} tr[\mathbf{I} - \mathbf{S}^{-1}\boldsymbol{\Sigma}(\hat{\boldsymbol{\Omega}})]^2,$$

where $tr(.)$ is the trace of the matrix, which is the sum of the diagonal elements in the matrix.

When one chooses $\mathbf{W}^{-1} = \hat{\boldsymbol{\Sigma}}(\boldsymbol{\Omega})^{-1}$, another discrepancy function based on Maximum Likelihood estimation is obtained as

$$F_{ML} = \log |\boldsymbol{\Sigma}_s(\hat{\boldsymbol{\Omega}})| + tr[\mathbf{S}\boldsymbol{\Sigma}_s^{-1}(\hat{\boldsymbol{\Omega}})] - \log |\mathbf{S}| - p,$$

where $\mathbf{S}$ is the sample covariance matrix, $p$ is the total number of observed variables, and |.| denotes the determinant of a matrix.

When the assumption of multivariate normality is tenable, the GLS and the ML fit functions have identical asymptotical properties. That is, they are asymptotically equivalent and also asymptotically efficient (Browne, 1974). As noted by Kaplan (2000,

p. 30), both GLS and ML functions are scale invariant and scale free under general conditions.

Becker and Olkin (2009, manuscript) also proved that a maximum likelihood estimator of the pooled correlation matrix, which is based on the asymptotic joint distribution of the correlations, is equal to the general least squares solution. Thus, at stage 1 of MASEM, the estimates of the pooled correlation matrix from both Becker's (1992) method and Cheung and Chan's (2005) should presumably lead to equivalent results. To estimate the model parameters in the modeling step (step 2 of TSSEM), Cheung and Chan (2005) used ADF as the estimation method, and it is expected that the standard errors of this method may differ from those based on the GLS method and the ML estimation method. To compare the performance of these multivariate methods, simulation studies are conducted. The design and results of the simulation studies are introduced in the next chapter.

# CHAPTER IV

# SIMULATION STUDY

To compare the empirical performance of the multivariate methods for MASEM, in this chapter I describe a Monte Carlo study to examine the alternative methods. Specifically, at the first step, the original GLS method (Becker, 1992), the modified GLS method (Becker & Fahrbach, 1994; Hafdahl, 2007), and the TSSEM method (Cheung & Chan, 2005) were applied for synthesizing the correlation matrices. At the second step, the resulting pooled correlation matrices were then subject to the second stage analysis of the TSSEM approach (i.e., path analysis using SEM software LISREL with ADF as the estimation method, proposed by Cheung, 2002), and also path modeling using Becker's (1992) direct computation method. In the first simulation study, homogeneous correlation matrices were generated based on a fixed-effects model. Then the first and second steps of the MASEM were analyzed by the three methods mentioned above. In the step one analyses, Type I error rates were examined for each approach. At the second step, parameter estimates and the standard errors of the parameter estimates were all examined for each method. The second simulation study applied a random-effects model to generate sample correlation matrices. Specifically, one element in the population correlation matrix was assumed random, thus the sample correlation matrices were generated from the 'random' parameters. In other words, the population correlation matrices for generating the sample correlations were assumed random samples from a 'super/hyper' population correlation matrix. Since both fixed-effects data and random-effects data were generated, in the first step of synthesizing correlation matrices, it was possible to cross the fixed-effects versus random-effects data and analyses. In this project, I examined three of the four combinations. Specifically, the analyses were fixed-effects analysis of fixed-effects data, fixed-effects analysis of random-effects data, and random-effects analysis of random-effects data. Since the TSSEM approach handles only fixed-effects analysis, the random-effects analysis of random-effects data was only

conducted with the two GLS approaches. At the second step, parameter estimates and the standard errors of the parameter estimates were also investigated.

## Simulation Conditions

Study 1: Fixed-effects Data

First, a fixed-effects model was applied in generating the sample correlation matrices. That is, I assumed that the sample correlation matrices were homogeneous and arose from one common population correlation matrix. The TSSEM and the GLS methods were respectively used for synthesizing the correlation matrices. With the TSSEM method, Maximum Likelihood Estimation (MLE) was used in the first step for parameter estimation and model evaluation as described in Cheung and Chan (2005). With the GLS methods, the multivariate $Q$ test in Equation 2.3 was used to test the homogeneity of the correlation matrices. In addition to investigating the original GLS method, one modification of the GLS method was also applied: instead of using the sample correlations, to compute the asymptotic variance-covariance matrix, the sample size weighted correlation coefficients were used in place of the population correlation coefficients (e.g., Becker & Fahrbach, 1994; Hafdahl, 2007). In synthesizing the correlation matrices at the first stage, all three approaches, namely, the TSSEM approach, the original GLS approach (GLS$r$), and the modified GLS approach (GLS$m$), were examined. At the second stage, two procedures were applied for the path analysis: 1) the TSSEM stage 2 analyses for path modeling; 2) GLS direct computation approach for both GLS$r$ and GLS$m$ at the first step. The parameter estimates and the standard errors of the parameter estimates were compared.

Data Generation

SAS/IML (SAS Institute, 1995) was used to generate correlation matrices following the data structure in Figure 3.1. LISREL 8.8 (LISREL, Jöreskög and Sorbom, 2000) was used for stage 1 and stage 2 analyses of the TSSEM approach and path

36

modeling with the Maximum Likelihood (ML) estimation method. Cheung (2007) developed a LISREL syntax generator for the two-stage analysis of MASEM and it was used in the simulation studies for getting the TSSEM results. Specifically, SAS was used as the program to call and run the TSSEM programs in the simulations (using method introduced in Gagne & Furlow, 2009). SAS/IML (SAS Institute, 1995) was used for the GLS methods of synthesizing correlation matrices and for direct computation of the path coefficients and standard errors. For comparison purposes, I used the same path models and simulation conditions as those chosen by Cheung (2002) for the fixed-effects-model data.

*Population correlation matrix ($\boldsymbol{\rho}$)*

A correlation matrix for $p = 4$ variables, with $p^* = p(p-1)/2 = 6$ correlation coefficients was created for each study with the population correlation matrix defined as

$$\boldsymbol{\rho} = \begin{matrix} & \begin{matrix} Y_1 & Y_2 & X_1 & X_2 \end{matrix} \\ \begin{matrix} Y_1 \\ Y_2 \\ X_1 \\ X_2 \end{matrix} & \begin{bmatrix} 1 & & & \\ .65 & 1 & & \\ .50 & .55 & 1 & \\ .58 & .35 & .20 & 1 \end{bmatrix} \end{matrix}. \tag{4.1}$$

In the first simulation study where the correlation matrices were homogeneous, this population correlation matrix served as the common population correlation matrix. It was built to produce the example path model shown in Figure 4.1. The population matrices presented in the example were defined as follows (also see Cheung, 2002, p. 71)

$$\boldsymbol{\Phi} = \begin{bmatrix} 1.0 & .2 \\ .2 & 1.0 \end{bmatrix}, \boldsymbol{\Gamma} = \begin{bmatrix} .4 & .5 \\ .3 & 0 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 0 \\ .5 & 0 \end{bmatrix}, \text{ and } \boldsymbol{\Psi} = \begin{bmatrix} .51 & 0 \\ 0 & .51 \end{bmatrix},$$

where $\boldsymbol{\Phi}, \boldsymbol{\Gamma}, \mathbf{B}, \boldsymbol{\Psi}$ corresponded to the matrices defined in Equation 3.1. Specifically, $\boldsymbol{\Phi}$ was the factor covariance matrix of the two independent variables; $\boldsymbol{\Gamma}$ was the matrix containing the path coefficients from the independent variables ($X_1$ and $X_2$) to the dependent variables ($Y_1$ and $Y_2$); $\mathbf{B}$ was the matrix containing the path coefficients from the dependent variable $Y_1$ to the dependent variable $Y_2$, and $\boldsymbol{\Psi}$ was the measurement error

matrix, respectively. Since there was no direct path from $X_2$ to $Y_2$, this model was an over-identified model with $df = 1$.

The path model with the population values is shown in Figure 4.1.
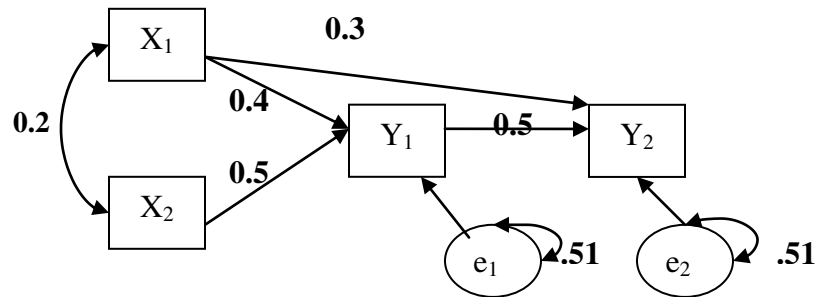


Figure 4.1 Example path analysis model 2 with the population values defined

First, to create correlation matrices based on a fixed-effects model, the population correlation matrix $\boldsymbol{\rho}$ was used. Multivariate normal distributed data were generated based on the population correlation coefficients in $\boldsymbol{\rho}$. SAS/IML functions RANNOR and ROOT (Cholesky decomposition) were used in generating the observed correlation matrix from each of the simulated primary studies, that is, after the observations were generated from a four-variate normal distribution with $\boldsymbol{\rho}$ being the specified population correlation matrix. Pearson correlation coefficients in $\mathbf{r}_i$ were computed from these observations to yield the fixed-effects data. The correlation coefficients computed from the raw data were then used for the subsequent analyses.

The simulation conditions considered were as described next (also see these simulation conditions in Cheung & Chan, 2005).

*Number of studies (k)*

Three values were chosen for the number of studies $k$: 5, 10, and 15. Cheung and Chan (2005, p. 49) noted that although the number of studies in real MASEM papers may be more than the number of studies listed above, the complexity of

the cross-grouop constrains in the Stage 1 of TSSEM as well as the inversion of the weight matrix with the ADF estimation at the Stage 2 analysis , computational challenges will arise in a simulation study with too many sample studies. Thus, Cheung (2002) and Cheung and Chan (2005) chose these numbers of studies to balance between the computation time and the generalizability of the simulation studies.

*Sample sizes within each study (n)*

Five levels of sample size per study were considered: 50, 100, 200, 500, and 1000.

Here $n = 50$ was used as the condition for a small sample and $n = 1000$ as the large sample condition (Cheung & Chan, 2005, p. 49).

Paxton, Curran, Bollen, Kirby, and Chen (2001) noted that 500 replications are usually large enough to give accurate statistical estimates in implementing Monte Carlo experiments in SEM. Thus, 500 replications were generated in these simulations. For simplicity, studies within each simulated meta-analysis had equal sample sizes and no missing correlations.

The $k \times n = 3 \times 5 = 15$ conditions are run for GLS*r*, GLS*m* and TSSEM methods of synthesizing *r* matrices and the resulting pooled correlation matrices were subject to path analysis using the model in Figure 4.1.

Study 2: Random-effects Data

Second, I used a random-effects model to generate the sample correlation matrices. That is, I assumed that the correlation matrices were from an average population correlation matrix rather than a single population matrix. To create correlation matrices based on a random-effects model, the population correlation matrix $\boldsymbol{\rho}$ as displayed in Equation 4.1 was treated as an 'average' population correlation matrix. One element in the correlation matrix $\boldsymbol{\rho}$ was generated as random. In this study, the correlation between $X_1$ and $Y_2$ was treated as random with mean 0.55 and standard deviation of 0.1. The 0.1 standard deviation is equivalent to a 0.01 variance of this correlation coefficient. Practically, this seems a small variance component and may be

prone to computational boundary conditions. However, bigger variances will result in inadmissible values of correlation coefficients, for example, values that exceed 1. The SAS/IML functions RANNOR and ROOT (Cholesky decomposition) were used in generating the observed correlation matrix from each of the simulated primary studies. In each replication, the observations were generated from a four-variate normal distribution with $\rho$ being the specified 'average' population correlation matrix, that is, the population correlation matrix used to generate the sample observations was assumed to be a random sample from the population correlation matrix with the specified mean and variance. At the path modeling step, the pooled correlation matrix was used as input to fit the path model in Figure 4.1 using TSSEM, GLS$r$ and GLS$m$ methods. The estimates of the path coefficients and standard errors were examined to understand how the first-step model changes affected the path-model parameter estimates at the second stage.

In the second study, first the fixed-effects analyses for synthesizing correlation matrices were run on the random-effects data. Specifically, the error variance considered in synthesizing the sample correlations was only the sampling error. As in the fixed-effects analyses of fixed-effects data situation, altogether $k \times n = 3 \times 5 = 15$ conditions were run for each of the methods (TSSEM, GLS$r$ and GLS$m$). Second, random-effects analyses were conducted on the random-effects data, where the between-studies variance was estimated and added to the sampling variance component. Then the estimates of the pooled correlation coefficients were obtained. Remember that the TSSEM approach would only be able to deal with fixied-effects analyses, thus the random-effects analyses of random-effects data were conducted only for the GLS methods. Also $k \times n = 3 \times 5 = 15$ conditions were run for each of the GLS methods (GLS$r$ and GLS$m$) in both steps.

Criteria for Parameter Estimates

*Relative Percentage Bias:*

The first criterion was the relative percentage bias of each parameter estimate and it was defined as

$$B(\hat{\Omega}) = \frac{\overline{\hat{\Omega}} - \Omega}{\Omega} \times 100\%,$$

where $\Omega$ was the population parameter, $\overline{\hat{\Omega}}$ was the average of the estimates of the parameters across all replications. Less than 5% bias was treated as acceptable and less than 2% bias was treated as good (Hoogland & Boomsma, 1998).

*Relative Percentage Bias of the Standard Error of the Parameter Estimate*

This measure of bias was defined as

$$B(S\overline{E}(\hat{\Omega})) = \frac{S\overline{E}(\hat{\Omega}) - SD(\hat{\Omega})}{SD(\hat{\Omega})} \times 100\%,$$

where $S\overline{E}(\hat{\Omega})$ was the mean of the estimated standard errors, and $SD(\hat{\Omega})$ was the empirical standard deviation of the parameter estimates across the replications. Less than 10% bias was treated as acceptable and less than 5% was treated as good (Hoogland & Boomsma, 1998). This index will show the closeness of the mean standard errors to the standard deviations of the parameter estimates.

*Standard Error of the Parameter Estimate:*

This index was defined as the mean value of the standard errors of the parameter estimates. It measures the precision of the parameter estimates. When the sample size increases, the standard errors of the parameter estimates should go down significantly.

*Confidence Interval of the Rejection Rates*

The confidence interval for the rejection rate with 500 replications can be approximated by the normal distribution. The 95% acceptance region for $\alpha = .05$ in the percentage metric was $5.0 \pm 1.96 \times \sqrt{(0.05) \times (0.95)/500} \times 100 = [3.09, 6.91]$.

**Results of the Simulation Studies**

Study 1: The Homogeneity Tests of the Correlation Matrices

*Rejection Rates of Each Method*

At the first stage of the TSSEM and GLS approaches, the homogeneity of the correlation matrices was tested. The GLS approaches used a multivariate $Q$ test and the TSSEM approach used the chi-square test built into the SEM software program with Multivariate Likelihood Estimation. The observed rejection rate of each approach at the first stage is shown in Table 4.1, with the expected rejection rate being 5% ($\alpha = 0.05$).

Table 4.1 shows that the original GLS method (GLS*r*) rejected the true model frequently when the sample sizes were small to medium ($n < 200$). The GLS*m* approach seemed to perform the best among the three methods, with rejection rates within the acceptable range for almost all conditions except the conditions $k = 10$, $n = 50$ and $k = 15$, $n = 200$. The TSSEM method controlled the Type I error rate pretty well at the larger sample sizes, but showed over-rejection at the smaller sample sizes. For example, for $k = 5$ and $k = 10$, the rejection rates for TSSEM and GLSr were all above 5% and out of the acceptable range when $n \leq 100$.

*Chi-square Test Statistics*

Table 4.2 shows the means and standard deviations of the chi-square test statistics from each approach and all conditions. For the GLS approaches, the chi-square test statistics were the multivariate $Q$ statistics calculated using Equation 2.7. For the TSSEM approach, the chi-square test statistics were the model fit test statistics using the maximum likelihood estimation at stage 1. For all three methods, the degrees of freedom of the test statistics were equal to $(k-1)p*$. Since I modeled no missing variables in each simulated condition, the expected degree of freedoms calculated were: $k = 5$, $df = 6(5-1)$ $= 24$; $k = 10$, $df = 6(10-1) = 54$; and $k = 15$, $df = 6(15-1) = 84$. As we know, the theoretical mean of the chi-square statistic equals its degrees of freedom. The expected standard deviation of the chi-square test statistic equals $\sqrt{2df}$ (cf. Freund, 1992). Thus, the expected standard deviations are: $k = 5$, $SD = 6.93$; $k = 10$, $SD = 10.39$; $k = 15$, $SD = 12.96$.

Figure 4.2 shows the means of the test statistics for each method. The test statistics for the GLS*r* method were generally positively biased. The bias was significant when the sample size per study was small ($n = 50$). When the sample sizes increased, the bias of the GLS*r* method decreased notably. The GLS*m* method performed much better

than the original GLS method, with the bias of the test statistics being minimal at all conditions. The TSSEM also performed well in all conditions except when $k = 15$ and $n = 1000$, where the test statistics tended to deviate from the expected means and the standard deviations exceeded the theoretical values. This is consistent with the idea of a long tail in the graph (see p. 24). Overall, the test statistics and the standard deviations of both the GLS$m$ and TSSEM approaches converged to the theoretical means when the sample size increased, with GLS$m$ being the most consistent among all three methods.

Study 1: Parameter Estimates and Standard Errors

At the second step, the pooled correlation matrices were subjected to path modeling using the example shown in Figure 4.1. The GLS direct computation methods produced only the estimates of the path coefficients and their standard errors. The TSSEM methods gave parameter estimates and standard errors as well as the overall model fit indices. Only parameter estimates and their standard errors were compared here.

*Relative Percentage Bias of the Path Coefficients*

The relative percentage bias values for the model parameter estimates were summarized for each of the four path coefficients ($\beta_{21}, \gamma_{21}, \gamma_{11}, \gamma_{12}$). Results showed that both the TSSEM method and the modified GLS method (i.e., the one with modified estimator of the asymptotic covariance matrix) yielded unbiased estimates of the path coefficients. Using 5% as an acceptable percentage for the relative bias (Hoogland & Boomsma, 1998), the path coefficients from these two methods were all unbiased, with relative percentages being within 2%. Table 4.3 shows the relative percentage bias of the parameter estimates obtained by different methods. The path coefficient estimates of the original GLS method were biased when the sample sizes were small ($n \leq 100$). For example, when $n = 50$, the bias was 5.39% for $\hat{\gamma}_{11}$ in the 10-studies condition. The relative percentage bias for $\hat{\beta}_{21}$ was 6.68% and for $\hat{\gamma}_{11}$ 6.34% in the 15-studies condition. This may because of the poor estimate of the asymptotic covariance matrix of correlation coefficients under small sample sizes (Becker & Wu, manuscript). When the sample size

increased, the path coefficients estimates computed with the original GLS method became less biased. Of all the three methods, the TSSEM approach resulted in the least relative bias, with all biases within 2%. In summary, both the TSSEM method and the modified GLS method estimated the path coefficients accurately, while the original GLS method gave less biased estimates of the path coefficients only when the sample size was larger than 100.

*Relative Percentage Bias of the Standard Errors of the Path Coefficients*

Table 4.4 summarizes the relative percentage bias of the standard errors of the path coefficient estimates from each method. Using 10% as an acceptable percentage for the standard error bias, the standard errors for path coefficient estimates from the TSSEM method were all biased for almost all conditions, with the highest percentages being around 30% and above. All the biases were positive, which means the standard errors were generally overestimated with the TSSEM method. This will affect confidence intervals and hypothesis tests of the path coefficients. In particular, the confidence intervals constructed using the TSSEM estimates will be wider than their true values, and hypothesis testing of the path coefficients may lead to more rejection decisions. On the other hand, GLS methods showed much lower relative percentage bias of the standard-error estimates. For the original GLS method, the relative percentage bias was generally higher than 10% when the sample size was small ($n = 50$). However, when sample size increased, the bias of the standard-error estimates decreased noticeably. When sample sizes were larger than 200, the standard-error estimates of the original GLS method had acceptable bias. The modified GLS method gave unbiased estimates of the standard errors for almost all conditions.

*Standard errors of the path coefficients*

The standard errors of the parameter estimates from each method under the fixed-effects data were plotted in Figure 4.10. The standard error reflects the precision of the estimation of the parameters. Figure 4.10 shows that, for each path coefficient, the standard error estimates from the GLS$m$ method and GLS$r$ method were very close to each other. For path coefficients $\beta_{21}$ and $\gamma_{21}$, the standard errors of the TSSEM estimates

were higher than those based on GLS*r* and GLS*m* methods. For path coefficients $\gamma_{11}$ and $\gamma_{12}$, the three methods showed comparable mean standard errors for all conditions. For all methods, when the sample size increased, the standard error decreased. The values of the standard error estimates were very small, ranging from .009 to .066 for the 5-studies conditions, from .007 to .046 for the 10-studies conditions, and from .005 to .024 for the 15-studies conditions across the three methods. Lower values of standard-error estimates are associated with more precision. The results also indicated that the differences between the standard errors of all the methods were actually small.

Study 2: Homogeneity Test of Correlation Matrices --- Fixed-effects Analyses

In study 2, random-effects model data were generated. First, fixed-effects analyses were conducted on these random-effects data using each of the three methods. In real-world analyses, fixed-effects analyses were preferred and applied when researchers decide that they want to generalize the results to the population in hand (Hedges & Vevea, 1998). On the other hand, it is not uncommon that fixed-effects analyses may be applied to random-effects data, especially in initial assessment of the nature of the data. This simulation study mimicked this practice.

For the GLS methods, multivariate *Q* tests were examined. For the TSSEM method, the first stage model chi-square test statistic was examined. As aforementioned, one element in the 'super/hyper' population correlation matrices was assumed random, and its standard deviation was .1, which means that the variance for that element was .01. Consequently, the between-study variance may be very small and applying a fixed-effects analysis on these random-effects data may still be reasonable.

Study 2: Parameter Estimates and Standard Errors --- Fixed-effects Analyses

In the second study, fixed-effects analyses were conducted on the random-effects data. Parameter estimates from each method were summarized and compared. Specifically, for the path model shown in Figure 4.1, the four path coefficients and their standard errors were examined.

*Rejection Rates of Each Method*

Table 4.5 shows the rejection rates for each method at the .05 significance level. The results shown here were consistent with those in Table 4.1. For the GLS$r$ approach, the rejection rates were big when sample sizes were medium to small $(n \leq 200)$. The highest rejection rates were 23% ($k = 10$, $n = 50$) and 26.8% ($k = 15$, $n = 50$). The rejection rates dropped rapidly when the within-study sample size $n$ reached 200 and then got closer to the nominal percentage at larger sample sizes ($n = 500$ and $n = 1000$). The GLS$m$ approach showed nominal rejection rates at almost all conditions, except when $k = 10$, $n = 50$, and $k = 15$, $n = 200$. The TSSEM approach had over-rejections at smaller sample sizes as well. For example, when sample size was 50, the rejection rates for the TSSEM approach were 8.4%, 11.2% and 8.2% for $k$ equals 5, 10 and 15, respectively. When $n > 100$, the rejection rates turned nominal quickly. Comparing the rejection rates of these three methods, the conclusion was that the GLS$m$ approach seemed to control the error rates better than the other approaches. When the within-study sample size became large ($n \geq 200$), all methods showed comparably values for the rejection rates.

*Chi-square Test Statistics*

The means and standard deviations of the chi-square test statistics are shown in Table 4.6. As described in the fixed-effects data situation, the expected degrees of freedom for the three $k$ conditions were $(k - 1)*p$ and they were calculated as 24, 54, and 84, respectively. The expected standard deviations of these chi-square statistics were 6.93, 10.39, and 12.96, respectively. We can see that the means and standard deviations of the chi-square statistics were very similar to those in Table 4.2, with only slight changes, mostly to the second decimal place of these values. The conclusion was that the chi-square statistics obtained from each of the GLS$r$, GLS$m$ and TSSEM methods converged to their theoretical means when the within-study sample sizes increased, with the GLS$m$ method being the most consistent among all three methods.

*Relative Percentage Bias of the Path Coefficients:*

The relative percentage biases for the model parameter estimates were calculated for each of the four path coefficients $\beta_{21}, \gamma_{21}, \gamma_{11}, \gamma_{12}$. The resulting bias values for the estimates of the path coefficients appear in Table 4.7. As can be seen from the table, no

substantial bias was found in the estimation of the path coefficient $\gamma_{12}$ parameter across all three methods. The bias estimates for this parameter ranged from 0% to 3.93%, with all values being within the 5% criterion for acceptable bias (Hoogland & Boomsma, 1998). When $n = 50$, the GLS$r$ method gave biased estimates of the parameters $\beta_{21}$ (5.54% ~ 6.75%) and $\gamma_{11}$ (5.43% ~ 6.22%). However, when sample size increased, the bias estimates dropped significantly. It appeared that the smallest bias values were for the TSSEM and GLS$m$ approaches, while the most biased estimates were from GLS$r$ method, especially at the small sample sizes conditions $(n \leq 100)$ .The similar situation was found for the path coefficients $\beta_{21}, \gamma_{21}$, and $\gamma_{11}$, where both the TSSEM approach and the GLS$m$ approach gave bias estimates within the acceptable range. Specifically, the TSSEM approach produced parameter estimates within 2% bias which was considered very good using the criteria proposed by Hoogland and Boomsma (1998). Except for the $n = 50$ sample-size condition, the most biased estimates among the four parameters were for $\beta_{21}$ and $\gamma_{21}$ regardless of the approach used for estimation. Overall, the modified GLS approach (GLS$m$) and the TSSEM approach yielded unbiased estimates of the path coefficients across all simulation conditions and the parameter estimates from GLS$r$ approach turned less biased when the sample sizes increased $(n \geq 100)$.

*Relative Percentage Bias of the Standard Errors of the Path Coefficients*

As stated previously, in creating the random population correlation matrix, the correlation between variables $X_1$ and $Y_2$ was treated as random with mean of .55 and standard deviation of .1. Direct decomposition of the correlation coefficients showed that the most affected path coefficients of this change would be $\beta_{21}$ (the relationship between $Y_1$ and $Y_2$) and $\gamma_{21}$ (the relationship between $X_1$ and $X_2$) (see Appendix A).

Table 4.8 displays the relative percentage bias of the standard errors. As can be seen in this table, the standard errors of parameters $\beta_{21}$ and $\gamma_{21}$ were all underestimated and were well beyond the acceptable range of bias regardless of the method used, with the range of underestimation being from 22.89% to over 95.08%. The estimates of the standard errors of the other two parameters $\gamma_{11}$ and $\gamma_{12}$ were generally not biased with the modified GLS approach (GLS$m$), with all bias being within the acceptable 10%

(Hoogland & Boomsma, 1998). The same situation was found for the standard error estimates for $\hat{\gamma}_{11}$ and $\hat{\gamma}_{12}$ using the original GLS method, except for the condition with $n$ = 50. With the TSSEM approach, the standard errors estimates of the parameters $\gamma_{11}$ and $\gamma_{12}$ were almost all overestimated across all conditions. In summary, the standard errors were generally underestimated by the original GLS$r$ method. The underestimation was huge when the sample sizes were small ($n$ = 50) and decreased rapidly when the sample sizes increased. The GLS$m$ approach produced unbiased estimates for the standard errors. The modified GLS approach usually had the smallest relative percentage bias in standard errors, and usually the bias decreased when the sample sizes increased. The relative biases of the TSSEM approach were also relatively large and beyond the acceptable range. The standard errors of the estimates of parameters $\gamma_{11}$ and $\gamma_{12}$ were overestimated by the TSSEM method. In other words, the confidence intervals constructed using the TSSEM estimates would be wider than the true values.

*Mean Standard Errors of the Path Coefficients*

The mean standard errors of the path coefficients estimates were summarized and plotted in Figure 4.17. As can be seen from this graph, the standard errors of the parameters were very small for each path coefficient, ranging from .011 to .066 for the 5-studies conditions, from .007 to .047 for the 10-studies conditions and from .005 to .038 for the 15-studies conditions across the three approaches. The lower values indicated higher precision. The standard error estimates based on the GLS$r$ method and the GLS$m$ method were essentially identical across the conditions. For the standard error estimates of the path coefficients $\beta_{21}$ and $\gamma_{21}$, the TSSEM estimates were larger than those based on GLS$r$ and GLS$m$ methods. For the standard errors of the path coefficients $\gamma_{11}$ and $\gamma_{12}$, the three methods produced equivalent mean standard errors at all conditions. For all the methods, when the sample size increased, the standard error decreased. This graph also demonstrated that the standard-error estimates of all three approaches were very small and close to each other.

Study 2: Parameter Estimates and Standard Errors --- Random-effects Analyses

As aforementioned, random-effects analyses of the random-effects data were only conducted for the two GLS approaches. Both the original GLS method and the modified GLS method produced parameter estimates within the acceptable range of relative percentage of bias. Figure 4.18 shows the relative percentage bias of the parameters for the GLS$m$ approach. As can be seen from the graph, all the biases were within 5%, with the most biased parameter estimates among the four parameters being for the parameters $\gamma_{21}$ and $\beta_{21}$. The results were very similar to those in the fixed-effects analyses.

Figure 4.19 shows the relative percentage bias of the standard errors of the parameters. With the random-effect analyses, both between-study variation and sampling variation were considered in estimating the parameters. It turned out that the standard errors of the parameter estimates were all exceeding the acceptable range, with the bias values for the standard-error estimates of the parameters $\gamma_{21}$ and $\beta_{21}$ being the most extreme.

The random-effects analyses of the random-effects data using the simulation conditions in this dissertation produced parameter estimates and standard error estimates similar to those in the fixed-effects analyses.

# CHAPTER V

# REAL DATA APPLICATIONS

As a part of my comparison of the multivariate methods for meta-analytic path modeling, one example was used to illustrate the application of these approaches to real data analysis. Data for this example came from a meta-analysis by Craft, Magyar, Becker and Feltz (2003). This meta-analysis study was conducted to disentangle the relationship between three components of anxiety and athletic performance. These three components were cognitive anxiety, somatic anxiety, and self-confidence (Martens, Vealey, & Burton, 1990a). These aspects of anxiety are often measured using the instrument Competitive State Anxiety Inventory (CSAI-2; Martens, Burton, Vealey, Bump, & Smith, 1990b). In this paper, the authors examined the interdependence of the three subscales and their relationship with athletes' performance, as well as the ability of the subscales of the CSAI-2 to predict athletic performance. Moderator analyses were also conducted to explore the possible effects of some study characteristics (type of sport, skill type, athlete's skill level, and time of CSAI-2 administration). The authors hypothesized that cognitive anxiety and performance would have a negative relationship; the relationship between performance and somatic anxiety would be very small and close to zero; performance and self-confidence would have a positive relationship; and that cognitive anxiety and somatic anxiety would have a positive relationship, but each would have a negative relationship with self-confidence (Craft et al., 2003, p. 50,).

The paper included 29 studies of the relationship between state anxiety measured by the CSAI-2 and athletic performance, among which 69 independent samples were observed. Both published and unpublished documents were collected and included. The authors indicated that twenty studies were published and nine were unpublished. The correlations were calculated using the observed Pearson's $r$ or they were transformed from Cohen's $d$ (Hedges & Olkin, 1985). In the original study, Craft et al. (2003) applied the Fisher's $Z$ transformation to the correlations. The transformed correlations were then weighted by the inverse of their variances. Multivariate analyses under a fixed-effects

model were also conducted using methods proposed by Becker (1992) and Becker and Schram (1994). Exploratory modeling was used to examine the interrelationships among the subscale variables and to explore the potential moderator effects of the relationship between anxiety and performance. Standardized regressions (path analyses) were examined to show the relationships between the independent and dependent variables while controlling the effects of the other variables in the model. The authors reported results for several subgroups, specifically, groups by sports type, type skills, level of athlete and the time of CSAI-2 administration.

Method

To illustrate how the multivariate MASEM methods are applied using real data and to compare their performance, I first synthesized the Craft et al. (2003) sample correlation matrices using the GLS*r*, GLS*m*, and the TSSEM stage 1 methods. Homogeneity tests were conducted to test if the sample correlation matrices were coming from the same population correlation matrix. At the second step, I applied the GLS direct computation method using the pooled correlation matrix to yield the estimates of the path coefficients and standard errors. The second stage of TSSEM was used as well for the analysis of the path models to estimate the path coefficients (standardized regression coefficients) and the standard errors. Moreover, in the stage 2 analysis of TSSEM, I applied maximum likelihood estimation (MLE) for the estimation of the parameters and standard errors to see how it compares with the ADF estimation method with the original TSSEM approach. The path-analysis model used here followed the theoretical structure in the original paper (Craft et al., 2003). The three subscales -- cognitive anxiety, somatic anxiety and self-confidence were assumed predictors of the athlete's performance. In path-analysis terminology, the athlete's performance was treated as the endogenous variable and the three anxiety subscales were treated as exogenous variables. Figure 5.1 shows the path diagram for the example. Among the 69 independent sample correlation matrices, 20 complete correlation matrices were found and selected for the purpose of this dissertation. Since there were three subscales of the CSAI-2 instrument (independent variables) and one performance variable (dependent variable), altogether six distinct correlations were contained in each correlation matrix. Table 5.1 shows the sample size

51

of each sample and the complete correlation matrices used in this analysis. As can be seen that the smallest sample size was 14 and the largest sample size was 199 in this study.

<div align="center">Results</div>

*Homogeneity Test of the Correlation Matrices*

At the first stage of synthesizing the correlation matrices, the GLS methods used the $Q$ statistic for testing the hypotheses. The multivariate $Q_E$ test was used to test the hypothesis that all of the variation in the observed correlation coefficients was due to random sampling variation. It is a chi-square test statistic with degrees of freedom equal to $k*(p-1)$, where $k$ is the number of studies (and $k$ equals 20 in this example), and $p = 4$ is the number of variables. With the TSSEM approach, maximum likelihood estimation was used in the first stage and the chi-square test statistic was used as a model test to see if the correlation matrices are homogeneous or not.

For the modified GLS approach, $Q_{GLSm}$ was 866.8, $df = 114$, p < .001. For the original GLS method, $Q_{GLSr} = 497.8$, $df = 114$, p < .001, whereas the chi-square test statistic of the TSSEM approach was 403.74, $df = 114$, p < .01. These results all showed that the correlation matrices were not homogeneous and the null hypothesis was rejected. The results suggested that the fixed-effects model may not be appropriate for pooling the correlation matrices. However, in real-world analyses, researchers do not always make the decision to use fixed-effects or random-effects analyses based solely on the homogeneity test. For example, fixed-effects analyses may still be applied to the data if the researcher only wanted to generalize the results to the populations in the studies at hand. In this project, multivariate fixed-effects analyses were first applied to estimate the synthesized correlation matrix, and then were used to get the second-step parameter estimates and standard errors.

*The Pooled Correlation Matrix and the Fitted Model*

Table 5.2 summarizes the pooled correlation matrices estimated from each of the approaches. As can be seen from the table, the pooled correlation matrices from the GLS*r* GLS*m* and TSSEM methods were very close in magnitude to each other.

The parameter estimates and their standard errors from each of the approaches were shown in Table 5.3. The estimates showed that the strongest relationship existed between the subscales self-confidence and performance. The estimated path coefficients from TSSEM and GLS$m$ were .408 and .516, respectively, whereas the standard error estimates were .033 and .028 for each of these approaches. The relationship was positive and significant, which means that more self-confidence in athletes leads to better performance. The weakest relationship was between the subscales somatic anxiety and performance, where the path coefficients estimates were -.019 by TSSEM and .069 by GLS$m$. The standard errors for this path coefficient were .033 by TSSEM and .031 by GLS$m$. We can see that estimates of parameter using these two methods were in opposite directions. Also, with the TSSEM approach, the test for this path coefficient was not significant ($z = -.572$, $p>.05$), however, the path coefficient was statistically significant with the GLS$m$ approach ($z = 2.21$, $p<.05$). For cognitive anxiety and performance, $\hat{\gamma}_{11} = .108$, $SE_{11} = .031$ by the TSSEM approach and the $z$-test statistic was 3.47, with $p < .05$; the estimate of this path coefficient by the GLS$m$ method had magnitude of .180, with standard error of .030. The $z$-test statistic was 6.04 and $p < .05$. In summary, both GLS and TSSEM showed cognitive anxiety was a significant predictor of athletes' performance, with the estimates from the GLS$m$ method demonstrating slightly stronger relationships than estimates from the TSSEM approach. The parameter estimates and standard-error estimates from the GLS$r$ method were also presented in Table 5.3. It turned out that this approach produced the strongest relationships between the predictors and the outcome, with all path coefficients being statistically significant using this approach.

At the second step of TSSEM, the maximum likelihood estimation was also applied instead of the ADF estimation in the original TSSEM method. It turned out that the parameter estimates were exactly the same as the ones estimated using ADF estimation, only with a slight difference in the estimates of the standard errors of the path coefficients. As noted in the methods section, the TSSEM approach uses the total sample size to fit SEM models in the second step. In this example, the total sample size was 1250, which was considered big enough for both ML and ADF estimation methods to perform well and produce equivalent parameter estimates.

In summary, the real example shows that overall the modified GLS method and the TSSEM method produced similar results in terms of both parameter estimates and standard errors estimates. The original GLS method produces parameter estimates slightly different from the other two approaches, resulting estimates showing stronger relationships among the variables. There were some differences in the parameter estimates and statistical tests which would lead to different conclusions about the overall relationships among the observed variables. Using maximum likelihood estimation at the second step of the TSSEM approach seemed to produce estimates very similar to those produced by the ADF estimation method. This particular real-data example mimicked the larger sample size and larger number of studies condition in the simulation study, with the within-study sample sizes being vary across studies. More empirical studies should be conducted to examine the generalizability of the simulation results to real data analyses.

# CHAPTER VI

# DISCUSSION

The purpose of this dissertation was to compare the multivariate methods for synthesizing correlation matrices and also different approaches for estimating path-analysis models using the synthesized correlation matrix. Specifically, the GLS approaches (Becker, 1992; Becker & Schram, 1994) and the TSSEM approach (Cheung, 2002; Cheung & Chan, 2005) were examined using both simulation studies and empirical research. The original GLS approach (Becker, 1992) is a multivariate approach of synthesizing correlation matrices, which consideres the dependence of the correlation coefficients by using the asymptotic covariance matrix of the correlation coefficients obtained by Olkin and Siotani (1976). Several researchers have conducted studies examining the statistical properties of this approach and the asymptotic correlation matrix used (e.g., Becker & Fahrbach, 1994; S. F. Cheung, 2000; Cheung & Chan, 2005; Furlow & Beretvas, 2005). These researchers all noted the poor performance of the asymptotic covariance matrix when the within-study sample sizes were small. In particular, Becker and Fahrbach (1994), S. F. Cheung (2000) and Furlow and Beretvas (2005) proposed some forms of modified estimators for the covariance matrix of the correlation coefficients. One of the modifications was to use the mean correlations in place of the observed individual correlation coefficients in the covariance matrix. This modification was applied in this dissertation and the modified GLS approach was examined along with the original GLS and TSSEM approaches.

## Conclusions and Suggestions

Results showed that the modified GLS approach performed much better than the original GLS approach not only in synthesizing the correlation matrices, but also in the estimation of the path coefficients and the standard errors of the path coefficients in the

second step of MASEM. Use of the modified GLS procedure avoided the problems found with the original GLS in the stage-one analysis, such as over-rejection of the true model in the hypothesis test of homogeneity in the first step of synthesizing correlation matrices. The modified GLS approach also produced less biased estimates of the parameters and their standard errors than the original GLS method. Compared to the TSSEM approach, the modified GLS procedure controlled the Type I error rates more consistently across all the simulation conditions. The chi-square test statistics from each of the three approaches converged to their theoretical means when sample sizes increased, with the modified GLS approach again being the most consistent compared to either the original GLS method or the TSSEM method.

The GLS direct computation procedures used in the second step provided parameter estimates and standard-error estimates equivalent with those from the SEM software and the TSSEM approach. In the second step of path analyses, both TSSEM and the modified GLS approach produced unbiased estimates of the path-coefficient parameters. The standard-error estimates from the modified GLS approach were the smallest among all three approaches, which means that this approach gives the most precise estimates of the parameters. However, the differences among the standard-error estimates of these methods were minimal, indicating that all three approaches estimated the standard errors accurately.

The performance of the three methods seemed to be similar for both fixed-effects data and random-effects data in this dissertation. Specifically, the fixed-effects analyses-of both the fixed-effects data and the random-effects data gave quite similar results in terms of the parameter estimates and the standard error estimates of the parameters for each of these approaches. However, because only one correlation coefficient was treated as random in generating the random-effects data, cautions should be given to the generalization of the results beyond this study.

Based on the simulation-study results and the real example shown in this dissertation, suggestions may be warranted for applied researchers conducting meta-analytic path-modeling analysis. First, the empirical performance of the multivariate methods discussed in this dissertation showed that both the modified GLS approach and the TSSEM approach were good at controlling the Type I error rate of the homogeneity

56

test when the within-study sample sizes were larger than 50. Either of them could be used in synthesizing correlation matrices. However, the original GLS method was not recommended when the within-study sample sizes were small to medium (less than 200), with which the rejection rate of the homogeneity test were out of the 95% confidence interval of the nominal rate of 5%. Second, in the second stage of fitting path models, both the direct computation approach and the TSSEM approach were appropriate for obtaining the parameter estimates and testing the significance of the parameters. The maximum likelihood estimation method which is the default estimation method for the path models may also be applied when the sample size is sufficient. The TSSEM approach has the advantage of providing different fit indices and overall model fit tests when the path model is not saturated.

## Limitations and Future Directions

This dissertation was designed to assess the performance of the multivariate approaches for MASEM, especially at the second stage of the path modeling process. The same fixed-effects simulation conditions were applied as those chosen by Cheung (2002) and Cheung and Chan (2005). For simplicity, sample sizes were assumed to be equal within a meta-analysis. However, as can be seen from the real example presented in Chapter 5, sample sizes varied greatly across studies. Unequal sample sizes are much more realistic in applied research. A more authentic data set could be generated with the sample sizes varying within meta-analyses. The largest number of studies examined in the simulation study was 15. Though it is not uncommon for researchers to conduct meta-analyses using less than 10 studies, it would also be reasonable to expand the number of studies and examine the performance of the methods under situations with larger numbers of studies. The parameter estimates would be more precise when the numbers of synthesized studies increase.

Several issues were not investigated in this dissertation. One is the usefulness of Fisher's (1921) *z* transformation. Researchers have found that using the Fisher-*z* transformation of the raw correlation coefficients may result in more accurate estimation of synthesized correlation coefficients and estimation of the structural model parameters

as well (Becker & Fahrbach, 1994; Cheung & Chan, 2005; Furlow & Beretvas, 2005; Hafdahl, 2001). In this project, only the raw correlation coefficients were studied. Future study may examine the Fisher-$z$ transformation and its impact on the second step of path analyses.

One other issue for synthesizing correlation matrices is the missing data problem. When examining the correlation matrices in real studies, researchers may not always be interested in the same set of variables in their studies. Thus it is not uncommon for there to be 'holes' in the examined correlation matrices. Pigott (1994) gave definitions for several patterns of missing data, such as missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). Researchers have conducted studies to examine the impact of different patterns of missing data on the synthesized correlation coefficients and the estimation of model parameters. Furlow and Beretvas (2005) concluded that MNAR data produced high levels of relative bias in correlations and model parameter estimates, and higher incorrect model rejection rates. It will be meaningful to eventually see how the different types of missing data affect the empirical performance of the GLS methods and the TSSEM methods at both the first and second stages of MASEM.

Another limitation was that this dissertation investigated only analyses based on the correct model that generated the data. No misspecified path models were examined. The path model examined in this project was very simple and no overall model fit indices were examined. In particular, the real data example showed a saturated path model. For situations where more complicated second-stage path models need to be defined and evaluated, the TSSEM approach provides fit indices at the second stage of the MASEM analyses, which is an advantage over the GLS direct computation methods, for which no overall fit index has been derived. The misspecification of the structural model presumably would impacts the parameter estimations and model tests. Future studies may examine misspecified models, and development of an overall fit index for the direct computation method would also be beneficial.

In addition, there were some limitations in generating the random-effects data for this dissertation. One correlation in the average population correlation matrix was assumed to be random and then the 'random' population correlation matrices were

generated. However, the random correlation's values were necessarily restricted, in that the variance of this element cannot be too big. When the variance component exceeds a certain value (usually a very small number like .01), the 'correlations' generated may exceed -1 or 1. Or it may lead to combinations or sets of correlation values that cannot exist together in any real correlation matrix. Because of this restriction, the resulting 'between-study' variation was very small in this dissertation. The small variation caused problems in estimating the between-study covariance matrix (e.g., subject to boundary restrictions, etc.). This low variance also led to there being only small differences between the random-effects data and the fixed-effects data, such that the empirical performance of the three methods was very similar in the random-effects data and fixed-effects data situations.
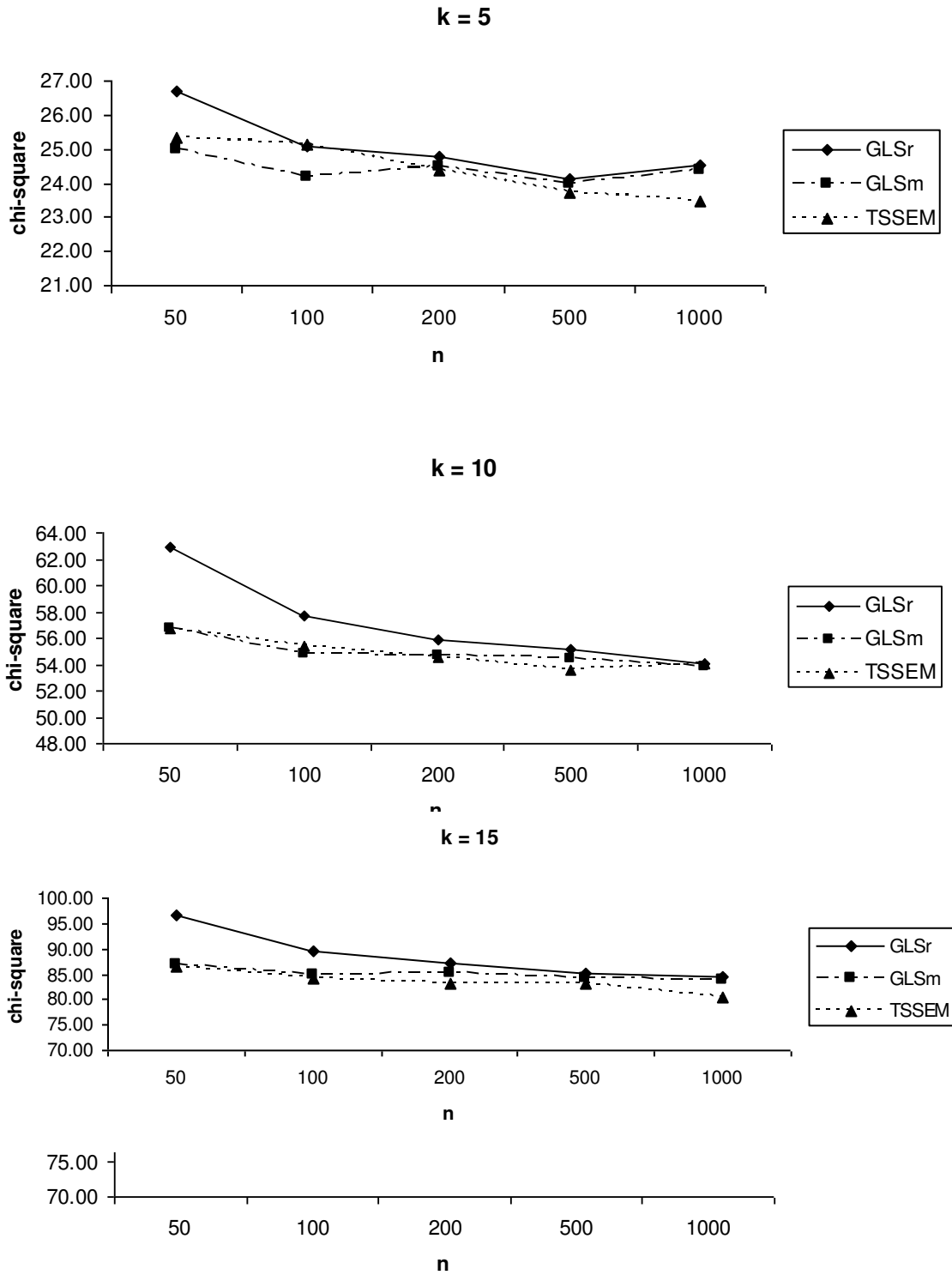
*Figure 4.2.* Means of chi-square statistics for the GLS*r*, GLS*m*, and TSSEM approaches in Study 1

**k = 5**



**k = 10**



**k = 15**
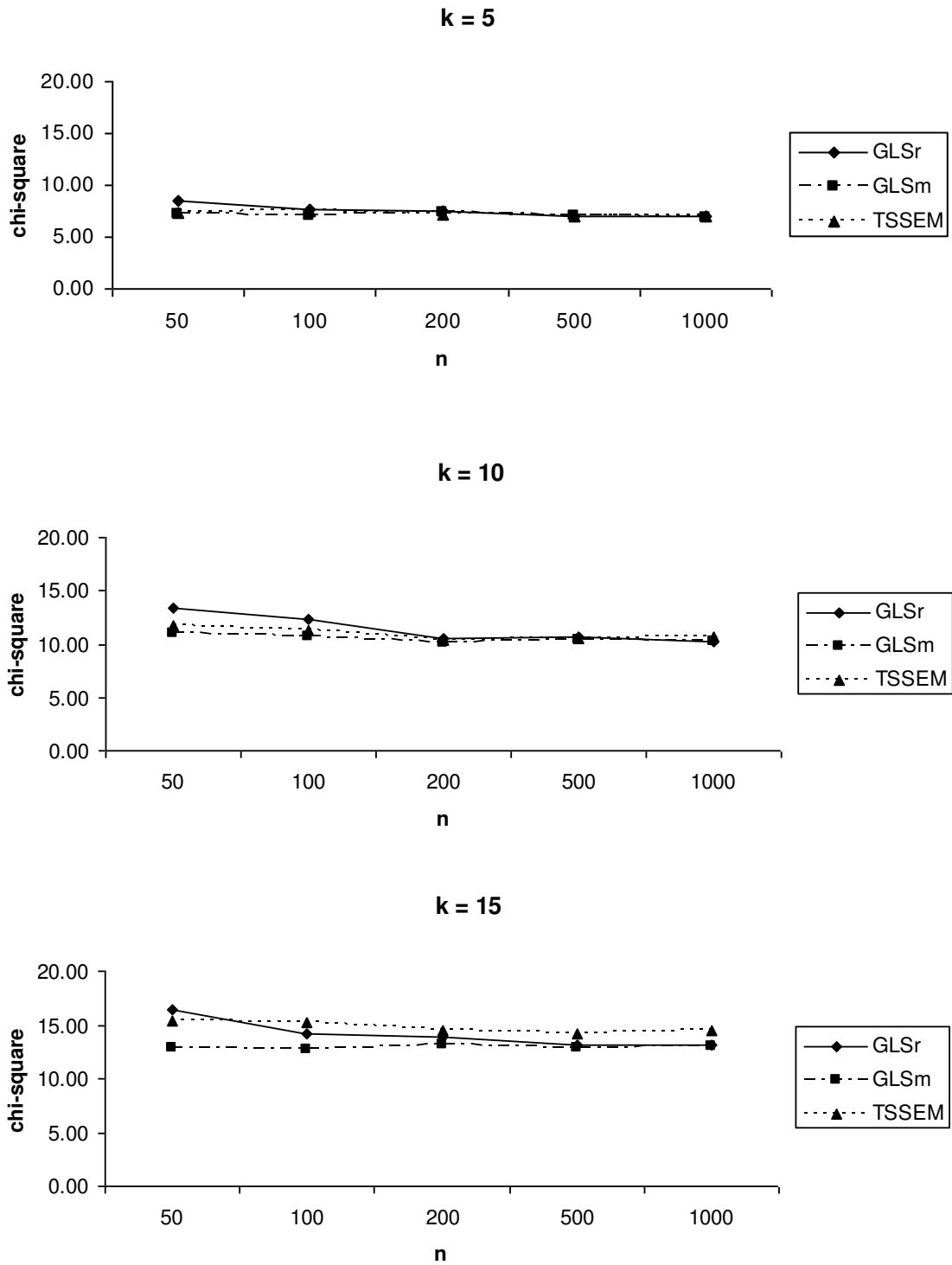


*Figure 4.3.* Standard deviations of chi-square statistics for the GLS*r*, GLS*m*, and TSSEM approaches in Study 1.
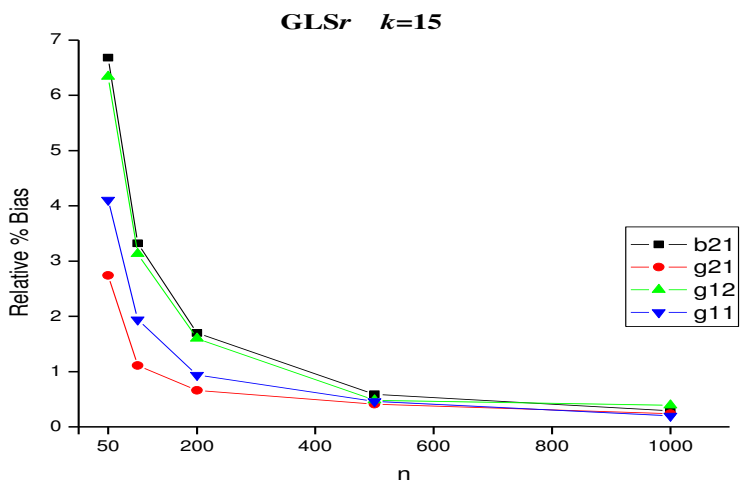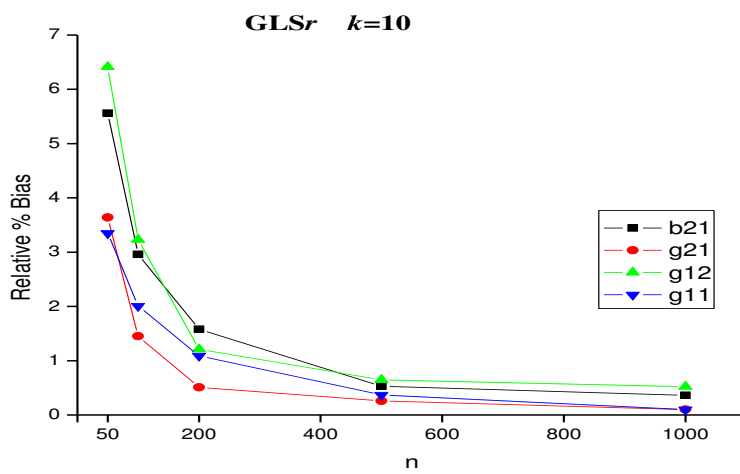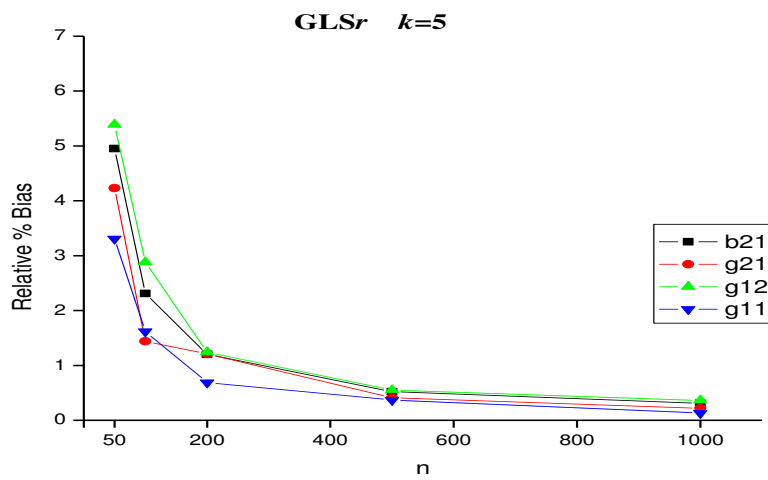
*Figure 4.4.* Relative % bias of the path coefficients for the GLSr approach in Study 1.
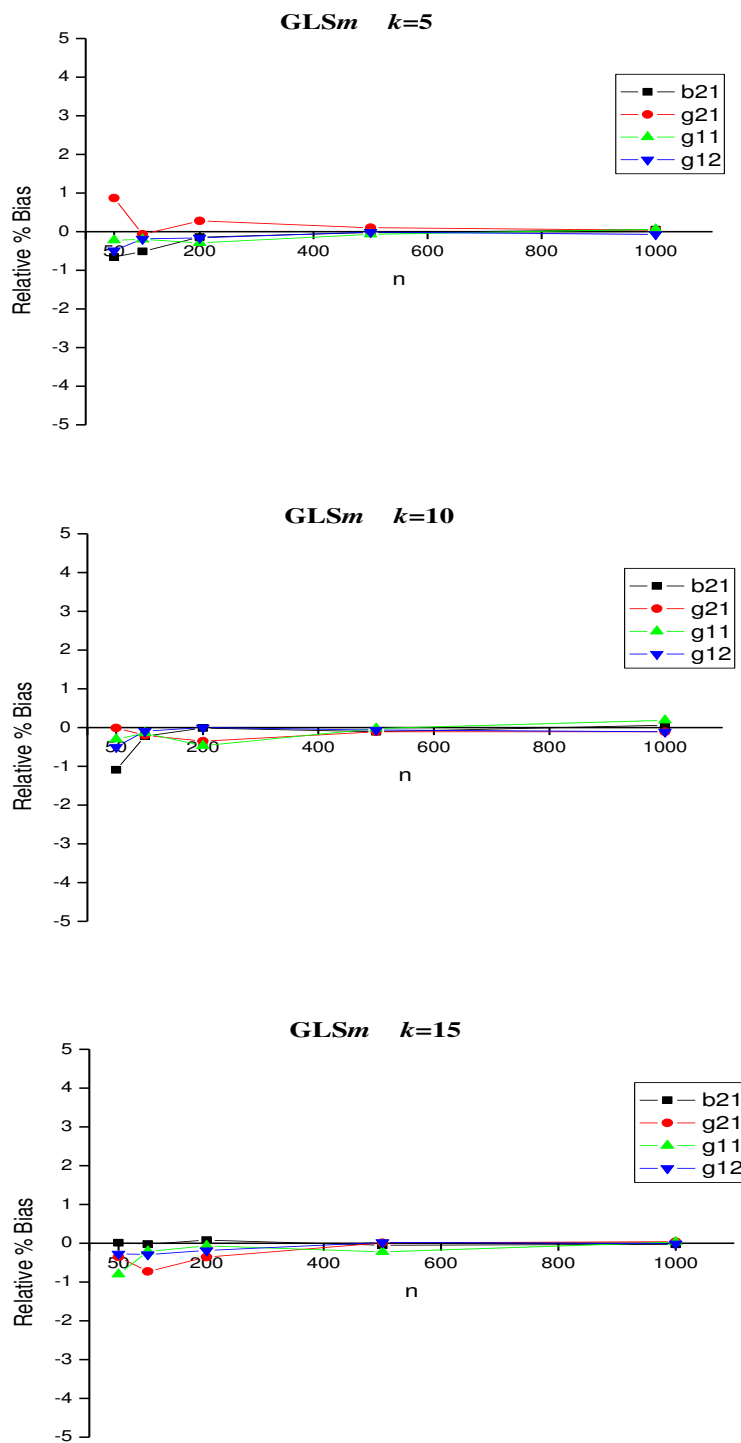
*Figure 4.5.* Relative % bias of the path coefficients for the GLS*m* approach in Study 1.
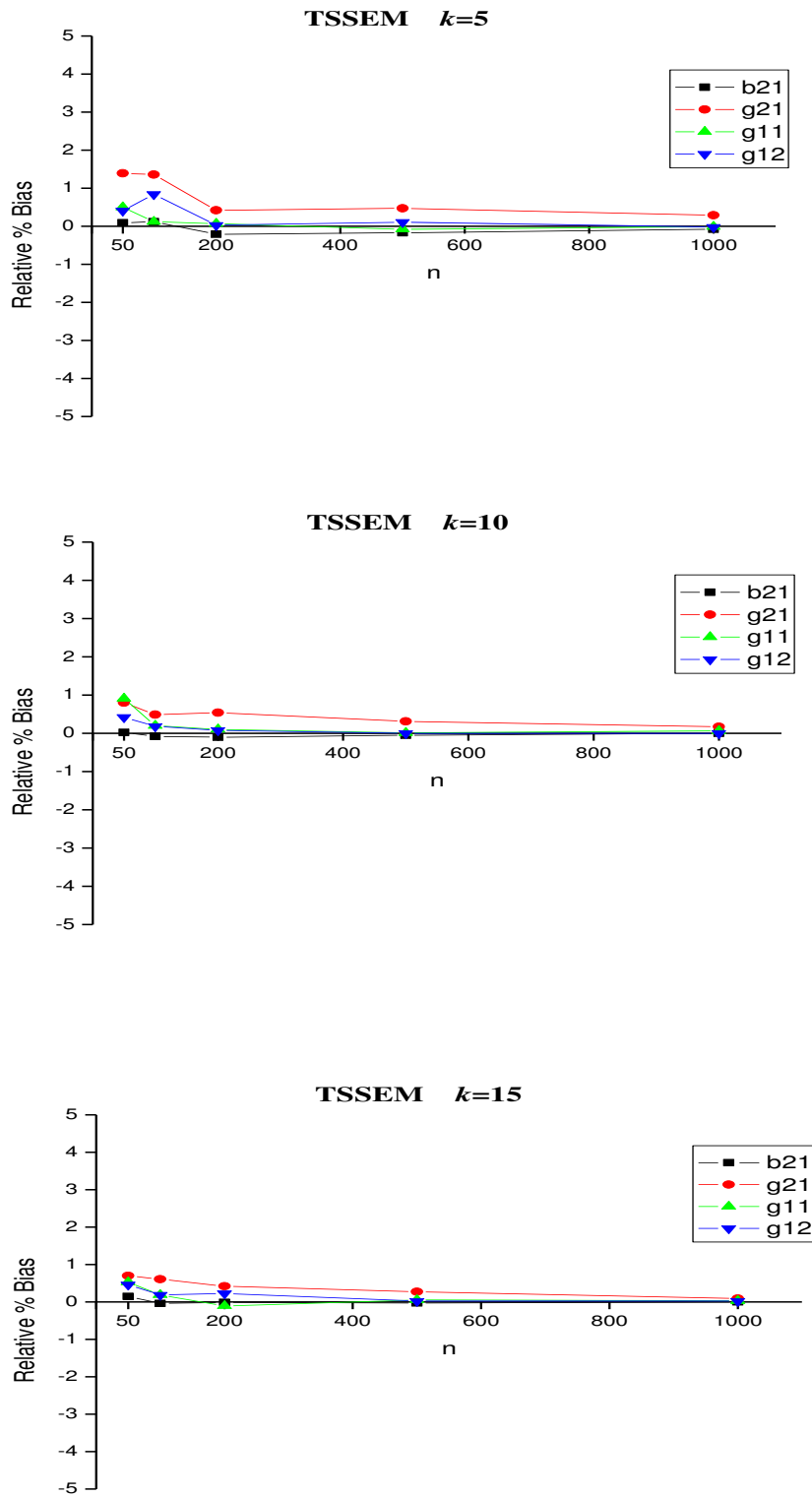
*Figure 4.6.* Relative % bias of the path coefficients for the TSSEM approach in Study1.
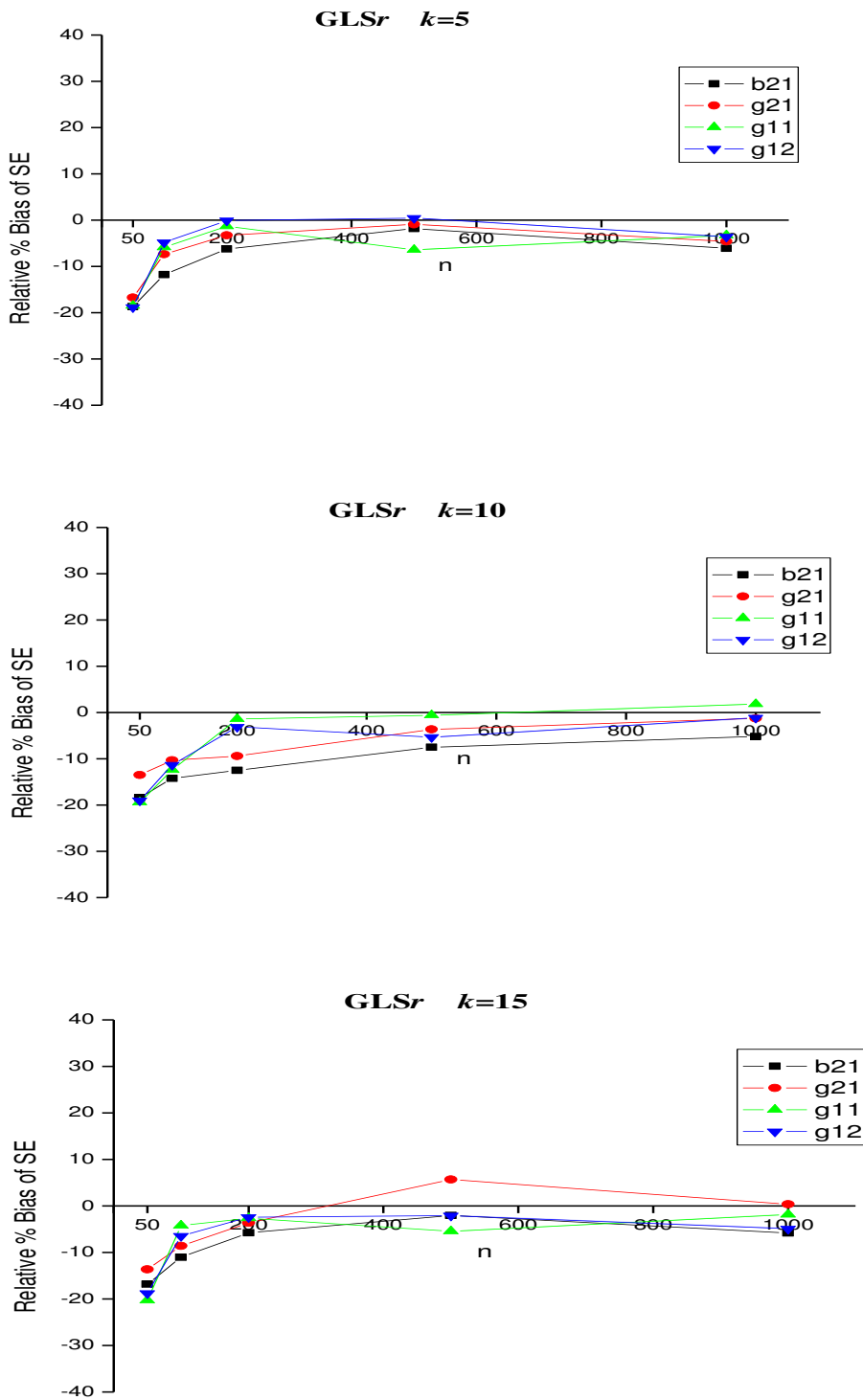
*Figure 4.7.* Relative % bias of the standard errors of parameter estimates for the GLS*r* approach in Study 1.
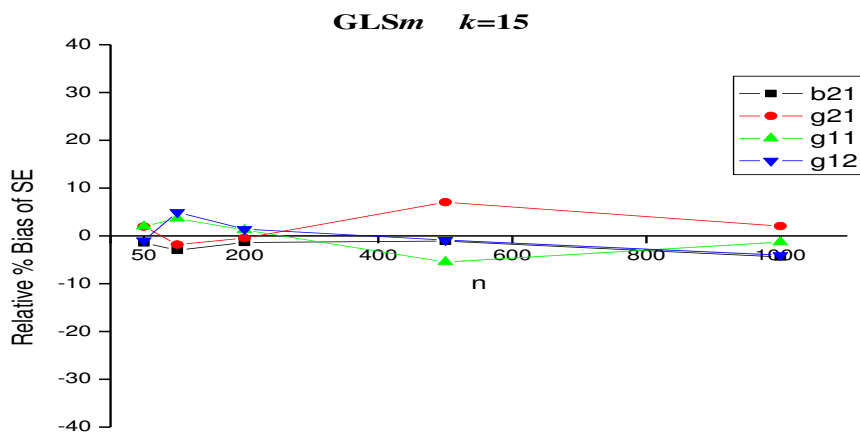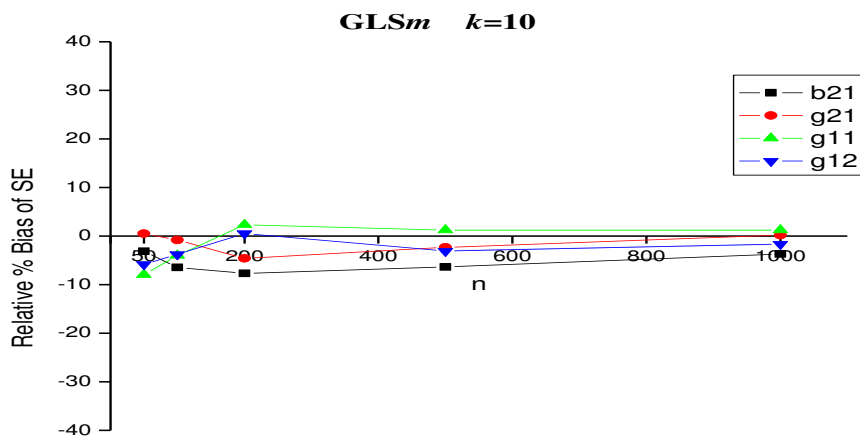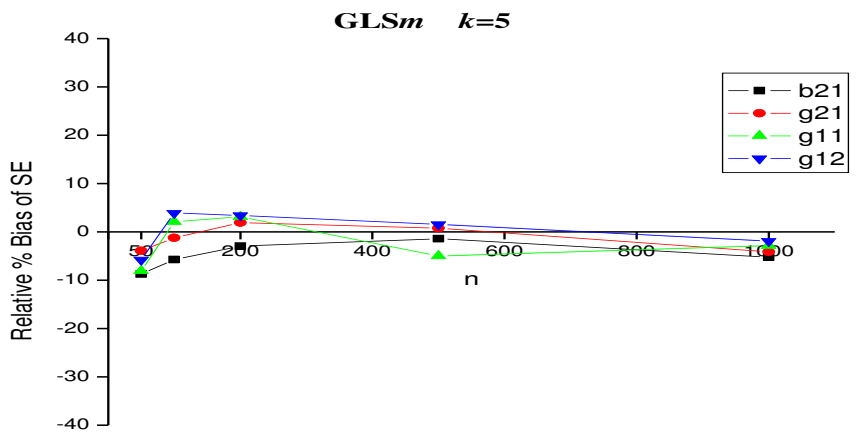
*Figure 4.8*. Relative % bias of the standard errors of parameter estimates for the GLS*m* approach in Study 1.
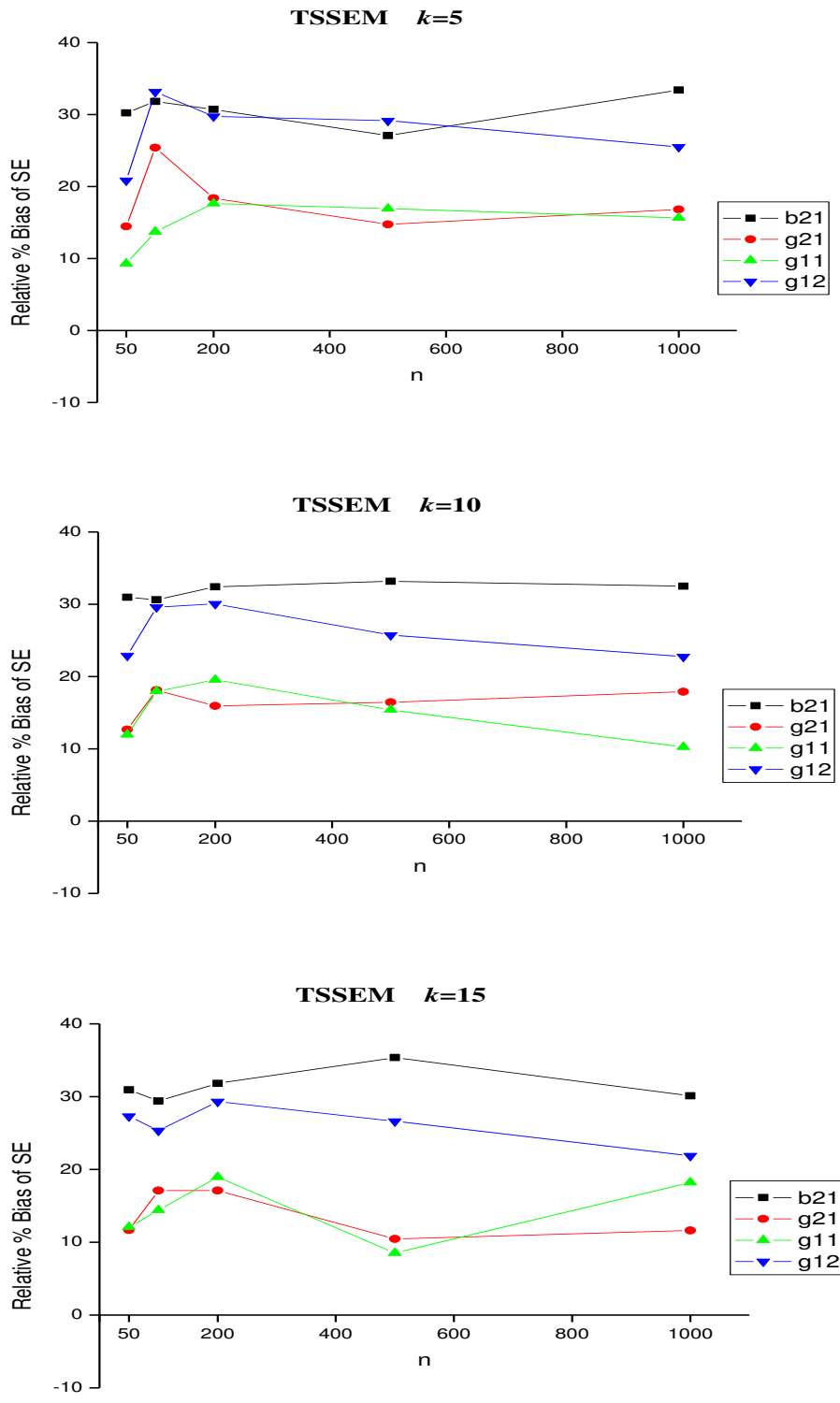
*Figure 4.9.* Relative % bias of the standard errors of parameter estimates for the TSSEM approach in Study 1.

Table 4.1

*Rejection Rate of Stage 1 of TSSEM, GLSr and GLSm in Study 1*

| k | n | TSSEM (%) | GLS*r* (%) | GLS*m* (%) |
|---|---|---|---|---|
| 5 | 50 | **8** | **12** | 6.2 |
|  | 100 | **8.4** | **7.8** | 6.0 |
|  | 200 | 6.2 | **7.6** | 5.2 |
|  | 500 | 4.2 | 5.2 | 4.8 |
|  | 1000 | 4.4 | 6.2 | 5.4 |
| 10 | 50 | **10.8** | **22.4** | **9.2** |
|  | 100 | **9.0** | **11.6** | 6.0 |
|  | 200 | 6.4 | 6.6 | 4.6 |
|  | 500 | 4.8 | 6.8 | 4.8 |
|  | 1000 | 5.2 | 4.6 | 4.0 |
| 15 | 50 | **8.2** | **26.8** | 6.6 |
|  | 100 | 5.2 | **12.2** | 5.8 |
|  | 200 | 3.8 | **9.4** | **7.4** |
|  | 500 | 4.4 | 6.6 | 6.0 |
|  | 1000 | 3.2 | 4.8 | 4.0 |

*Note.* the bold faces numbers are the percentages that are outside of

the 95% confidence interval of the rejection rate.

Table 4.2

*Chi-square Statistics and Their Standard Deviations in study 1*

| $k$ | $n$ | GLS*m* | | GLS*r* | | TSSEM | |
|---|---|---|---|---|---|---|---|
| | | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* |
| 5 | 50 | 24.99 | 7.20 | 26.70 | 8.55 | 25.36 | 7.40 |
| | 100 | 24.20 | 7.05 | 25.10 | 7.62 | 25.13 | 7.54 |
| | 200 | 24.49 | 7.26 | 24.77 | 7.46 | 24.36 | 7.23 |
| | 500 | 23.96 | 6.95 | 24.11 | 7.05 | 23.71 | 7.07 |
| | 1000 | 24.40 | 6.85 | 24.52 | 7.05 | 23.45 | 7.02 |
| 10 | 50 | 56.70 | 11.05 | 62.96 | 13.40 | 56.67 | 11.68 |
| | 100 | 54.77 | 10.75 | 57.64 | 12.30 | 55.36 | 11.31 |
| | 200 | 54.66 | 10.08 | 55.86 | 10.56 | 54.54 | 10.41 |
| | 500 | 54.45 | 10.32 | 55.15 | 10.73 | 53.61 | 10.56 |
| | 1000 | 53.79 | 10.23 | 54.11 | 10.29 | 54.01 | 10.62 |
| 15 | 50 | 86.78 | 12.87 | 96.60 | 16.39 | 86.39 | 15.30 |
| | 100 | 84.92 | 12.67 | 89.67 | 14.11 | 84.04 | 15.29 |
| | 200 | 85.02 | 13.19 | 87.23 | 13.90 | 83.26 | 14.46 |
| | 500 | 84.28 | 12.90 | 85.05 | 13.20 | 83.01 | 14.14 |
| | 1000 | 83.93 | 13.01 | 84.44 | 13.07 | 80.61 | 14.41 |

*Note.* Theoretical means and standard deviations are 24 and 6.93 for $k = 5$, 54 and 10.39 for $k = 10$ and 84 and 12.96 for $k = 15$.

Table 4.3

*Relative Percentage Bias of Path Coefficient Estimates in Study 1*

| | | k = 5 | | | | k = 10 | | | | k = 15 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *n* | Methods | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ |
| 50 | TSSEM | .09 | 1.39 | .50 | .41 | .02 | .80 | .91 | .15 | .70 | .53 | .46 | .15 |
| | GLS*r* | 4.95 | 4.23 | 5.39 | 3.31 | 5.56 | 3.64 | 6.41 | 3.35 | 6.68 | 2.74 | 6.34 | 4.11 |
| | GLS*m* | -0.66 | .87 | -.22 | -.49 | -1.09 | -.01 | -.30 | -.50 | .01 | -.35 | -.80 | -.27 |
| 100 | TSSEM | .12 | 1.36 | .12 | .84 | -.08 | .49 | .20 | -.04 | .61 | .19 | .19 | -.04 |
| | GLS*r* | 2.31 | 1.44 | 2.88 | 1.62 | 2.96 | 1.45 | 3.23 | 2.01 | 3.32 | 1.11 | 3.13 | 1.94 |
| | GLS*m* | -.51 | -.07 | -.20 | -.19 | -.22 | -.20 | -.14 | -.09 | -.02 | -.73 | -.21 | -.29 |
| 200 | TSSEM | -.21 | .42 | .07 | .03 | -.10 | .54 | .10 | -.01 | .42 | -.11 | .23 | -.01 |
| | GLS*r* | 1.20 | 1.21 | 1.24 | .69 | 1.58 | .51 | 1.21 | 1.09 | 1.70 | .66 | 1.60 | .94 |
| | GLS*m* | -.14 | .28 | -.29 | -.16 | -.01 | -.35 | -.47 | .00 | .08 | -.36 | -.07 | -.19 |
| 500 | TSSEM | -.17 | .47 | -.08 | .11 | -.05 | .31 | .01 | .00 | -.03 | .27 | .05 | .02 |
| | GLS*r* | .52 | .41 | .55 | .37 | .53 | .26 | .65 | .37 | .59 | .41 | .48 | .46 |
| | GLS*m* | -.02 | .10 | -.07 | -.01 | -.11 | -.11 | -.02 | -.06 | -.05 | .01 | -.22 | .02 |
| 1000 | TSSEM | -.08 | .29 | -.01 | -.02 | .00 | .17 | .06 | .00 | .00 | .09 | .03 | .02 |
| | GLS*r* | .31 | .22 | .36 | .13 | .36 | .10 | .52 | .10 | .29 | .24 | .39 | .20 |
| | GLS*m* | .05 | .04 | .06 | -.07 | .06 | -.11 | .19 | -.11 | -.03 | .04 | .03 | -.01 |

*Note.* GLS*r* represents the original GLS method which uses sample correlations as substitute for population correlations in the asymptotic covariance matrix; GLS*m* represents the modified estimator of the asymptotic covariance matrix using mean correlation coefficient

Table 4.4

*Relative Percentage Bias of the Mean Standard Errors of the Path Coefficients Estimate in Study1*

| | | $k = 5$ | | | | $k = 10$ | | | | $k = 15$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Methods | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ |
| 50 | TSSEM | 30.23 | 14.45 | 9.29 | 20.85 | 30.98 | 12.66 | 11.95 | 22.89 | 30.93 | 11.69 | 12.07 | 27.31 |
| | GLS$r$ | -18.69 | -16.71 | -18.49 | -18.83 | -18.43 | -13.50 | -19.39 | -19.01 | -16.81 | -13.64 | -20.33 | -18.83 |
| | GLS$m$ | -8.71 | -3.92 | -8.08 | -5.79 | -3.12 | .50 | -7.94 | -5.82 | -1.44 | 1.96 | 1.98 | -.95 |
| 100 | TSSEM | 31.81 | 25.40 | 13.72 | 33.15 | 30.62 | 18.13 | 17.96 | 29.61 | 29.39 | 17.09 | 14.40 | 25.33 |
| | GLS$r$ | -11.79 | -7.37 | -5.87 | -4.81 | -14.22 | -10.30 | -12.43 | -11.30 | -11.04 | -8.57 | -4.24 | -4.81 |
| | GLS$m$ | -5.74 | -1.24 | 2.07 | 3.93 | -6.47 | -0.83 | -3.99 | -3.77 | -3.01 | -1.84 | 3.63 | 4.94 |
| 200 | TSSEM | 30.72 | 18.38 | 17.65 | 29.75 | 32.42 | 15.93 | 19.54 | 30.07 | 31.81 | 17.09 | 18.94 | 29.32 |
| | GLS$r$ | -6.23 | -3.27 | -1.30 | -.11 | -12.51 | -9.40 | -1.40 | -3.08 | -5.76 | -3.62 | -2.62 | -.11 |
| | GLS$m$ | -2.95 | 1.93 | 3.13 | 3.41 | -7.67 | -4.57 | 2.33 | .55 | -1.34 | -0.47 | 1.24 | 1.43 |
| 500 | TSSEM | 27.06 | 14.75 | 16.92 | 29.14 | 33.19 | 16.45 | 15.40 | 25.75 | 35.37 | 10.45 | 8.53 | 26.63 |
| | GLS$r$ | -1.81 | -.90 | -6.40 | .44 | -7.50 | -3.63 | -.60 | -5.30 | -2.05 | 5.66 | -5.46 | 0.44 |
| | GLS$m$ | -1.39 | .79 | -4.99 | 1.55 | -6.36 | -2.34 | 1.23 | -3.08 | -1.08 | 7.06 | -5.47 | -0.88 |
| 1000 | TSSEM | 33.42 | 16.82 | 15.64 | 25.52 | 32.49 | 17.90 | 10.25 | 22.75 | 30.12 | 11.61 | 18.22 | 21.87 |
| | GLS$r$ | -6.11 | -4.61 | -3.29 | -3.61 | -5.20 | -1.25 | 1.81 | -1.12 | -5.81 | 0.38 | -1.85 | -3.61 |
| | GLS$m$ | -5.25 | -4.14 | -2.87 | -1.92 | -3.68 | .18 | 1.25 | -1.63 | -4.42 | 2.05 | -1.30 | -4.06 |

*Note.* GLS$r$ represents the original GLS method which uses sample correlations as substitute for population correlations in the asymptotic covariance matrix; GLS$m$ represents the modified estimator of the asymptotic covariance matrix using mean correlation coefficient
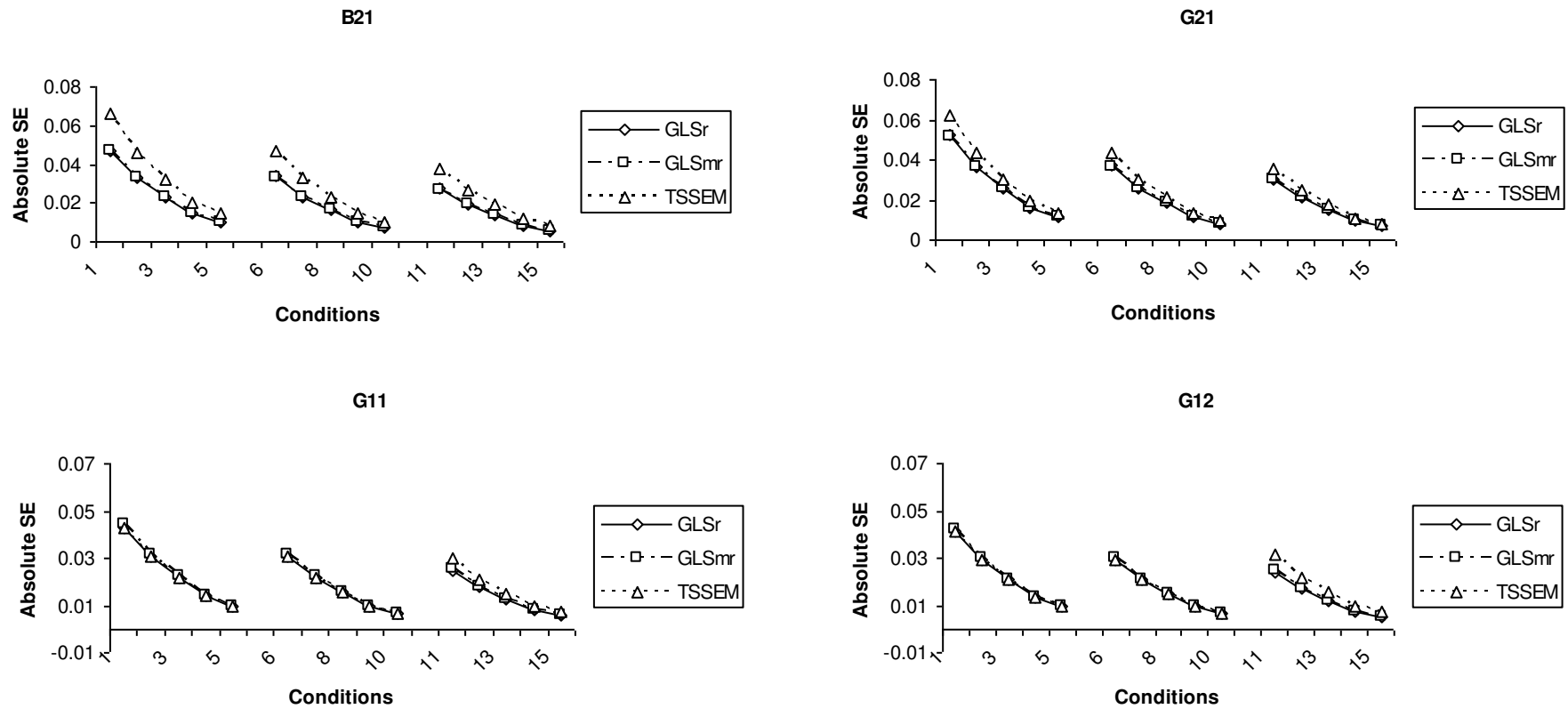
*Figure 4.10.* Mean standard errors of the path coefficients of each method in Study 1

Table 4.5

*Rejection Rates at Stage 1 of TSSEM, GLSr and GLSm in Study 2*

| k | n | TSSEM (%) | GLS*r* (%) | GLS*m* (%) |
|---|---|---|---|---|
| 5 | 50 | **8.6** | **12.2** | 5.8 |
| | 100 | **8.2** | **7.4** | 5.6 |
| | 200 | 6.4 | **7.4** | 5.6 |
| | 500 | 4.0 | 5.2 | 4.8 |
| | 1000 | 4.6 | 6.4 | 5.2 |
| 10 | 50 | **10.2** | **23** | **8.6** |
| | 100 | **9.0** | **11.4** | 6.0 |
| | 200 | 6.4 | 6.6 | 4.2 |
| | 500 | 5.0 | 6.8 | 5.2 |
| | 1000 | 4.8 | 4.8 | 4.6 |
| 15 | 50 | **8.2** | **26.8** | 6.8 |
| | 100 | 4.6 | **12.2** | 5.6 |
| | 200 | 4.2 | **9.4** | **7.2** |
| | 500 | 5.4 | 6.6 | 6.0 |
| | 1000 | 3.2 | 4.6 | 4.2 |

*Note:* The percentages that are out of the 95% confidence interval of the rejection rate are presented in bold face.

Table 4.6

*Chi-square Statistics and Their Standard Deviations in Study 2*

| k | n | GLS*m* | | GLS*r* | | TSSEM | |
|---|---|---|---|---|---|---|---|
| | | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* |
| 5 | 50 | 25.00 | 7.19 | 26.73 | 8.56 | 25.48 | 7.47 |
| | 100 | 24.18 | 7.06 | 25.10 | 7.62 | 25.12 | 7.57 |
| | 200 | 24.50 | 7.28 | 24.79 | 7.48 | 24.34 | 7.27 |
| | 500 | 24.00 | 7.01 | 24.13 | 7.09 | 23.73 | 7.08 |
| | 1000 | 24.43 | 6.88 | 24.53 | 7.06 | 23.46 | 7.00 |
| 10 | 50 | 56.71 | 11.02 | 56.71 | 13.35 | 56.54 | 11.89 |
| | 100 | 54.80 | 10.69 | 54.80 | 12.29 | 55.28 | 11.38 |
| | 200 | 54.67 | 10.06 | 54.67 | 10.57 | 54.54 | 10.41 |
| | 500 | 54.47 | 10.34 | 54.47 | 10.74 | 53.68 | 10.54 |
| | 1000 | 53.86 | 10.23 | 53.86 | 10.28 | 54.00 | 10.58 |
| 15 | 50 | 86.77 | 12.92 | 86.77 | 16.37 | 86.25 | 15.54 |
| | 100 | 84.98 | 12.64 | 84.98 | 14.08 | 84.03 | 15.34 |
| | 200 | 85.15 | 13.26 | 85.15 | 13.92 | 83.26 | 14.44 |
| | 500 | 84.25 | 12.99 | 84.25 | 13.26 | 82.95 | 14.18 |
| | 1000 | 83.98 | 12.89 | 83.98 | 13.00 | 80.86 | 14.20 |

Table 4.7

*Relative Percentage Bias of Path Coefficient Estimates in Study 2*

| | | $k = 5$ | | | | $k = 10$ | | | | $k = 15$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | Methods | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ |
| 50 | TSSEM | .63 | .16 | .61 | .40 | .64 | -.76 | .96 | .35 | .82 | -.88 | .50 | .46 |
| | GLS$r$ | 5.54 | 1.92 | 5.43 | 3.33 | 6.50 | 1.15 | 6.52 | 3.70 | 6.75 | 1.44 | 6.22 | 3.93 |
| | GLS$m$ | -.16 | -.86 | -.20 | -.52 | -.37 | -1.42 | .01 | -.59 | -.24 | -1.40 | -.45 | -.56 |
| 100 | TSSEM | .44 | -.86 | .69 | .16 | .47 | -1.03 | .24 | .16 | .47 | -.77 | .22 | .18 |
| | GLS$r$ | 2.23 | 1.41 | 2.96 | 1.59 | 2.59 | 2.42 | 2.79 | 1.77 | 2.87 | 2.25 | 3.15 | 1.91 |
| | GLS$m$ | -.38 | -.10 | -.28 | -.16 | -.48 | .51 | -0.50 | -.34 | -.48 | .69 | -.36 | -.22 |
| 200 | TSSEM | .34 | -1.09 | .14 | -.03 | .46 | -.97 | .15 | .07 | .57 | -1.08 | -.08 | .22 |
| | GLS$r$ | 1.99 | -1.51 | 1.28 | .69 | 2.31 | -1.68 | 1.40 | 1.09 | 2.21 | -1.50 | 1.47 | .96 |
| | GLS$m$ | .67 | -2.38 | -.23 | -.17 | .70 | -2.55 | -.20 | -.02 | .69 | -2.61 | -.16 | -.21 |
| 500 | TSSEM | .44 | -.98 | -.07 | .12 | .57 | -1.22 | .00 | .01 | .62 | -1.38 | .05 | -.02 |
| | GLS$r$ | 1.50 | -2.90 | .56 | .38 | 1.61 | -2.96 | .59 | .38 | 1.66 | -3.09 | .68 | .45 |
| | GLS$m$ | .98 | -3.21 | -.10 | -.02 | 1.02 | -3.26 | -.02 | -.09 | 1.03 | -3.43 | -.03 | .01 |
| 1000 | TSSEM | .55 | -1.23 | -.03 | -.01 | .57 | -1.29 | .05 | .00 | .59 | -1.51 | .05 | -.01 |
| | GLS$r$ | .94 | -1.87 | .35 | .14 | .88 | -1.96 | .32 | .20 | .95 | -1.98 | .48 | .13 |
| | GLS$m$ | .68 | -2.04 | .03 | -.08 | .58 | -2.13 | .00 | -.02 | .62 | -2.12 | .13 | -.08 |

*Note.* GLS$r$ represents the original GLS method which uses sample correlations as substitute for population correlations in the asymptotic covariance matrix; GLS$m$ represents the modified estimator of the asymptotic covariance matrix using mean correlation coefficient

Table 4.8

*Relative Percentage Bias of the Mean Standard Errors of the Path Coefficients Estimate in Study 2*

| $n$ | Methods | $k = 5$ | | | | $k = 10$ | | | | $k = 15$ | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ | $\beta_{21}$ | $\gamma_{21}$ | $\gamma_{11}$ | $\gamma_{12}$ |
| 50 | TSSEM | -22.89 | -56.70 | 10.20 | 20.90 | -39.70 | -68.76 | 12.25 | 21.54 | 28.69 | -49.78 | -74.60 | 13.07 |
| | GLSr | -49.13 | -65.07 | -18.68 | -18.90 | -60.52 | -74.17 | -20.26 | -21.05 | -67.36 | -78.94 | -19.45 | -17.94 |
| | GLSm | -42.71 | -63.85 | -8.64 | -5.80 | -56.14 | -74.35 | -3.90 | -4.61 | -63.01 | -78.34 | -6.56 | -1.88 |
| 100 | TSSEM | -39.40 | -68.80 | 14.33 | 24.44 | -55.71 | -77.75 | 19.12 | 30.59 | -63.64 | -81.73 | 14.95 | 25.46 |
| | GLSr | -56.10 | -72.80 | -5.20 | -4.45 | -68.15 | -80.46 | -4.68 | -4.96 | -73.56 | -84.03 | -8.85 | -10.61 |
| | GLSm | -55.71 | -73.25 | 1.57 | 3.32 | -65.68 | -80.21 | 2.98 | 5.66 | -72.08 | -83.82 | -1.10 | -1.18 |
| 200 | TSSEM | -55.62 | -77.76 | 16.03 | 24.39 | -68.09 | -84.29 | 19.08 | 29.47 | -73.47 | -87.10 | 18.66 | 29.32 |
| | GLSr | -65.42 | -79.80 | -1.42 | -0.06 | -75.08 | -85.72 | -5.74 | -4.72 | -79.80 | -88.25 | -6.91 | -5.55 |
| | GLSm | -64.54 | -79.72 | 2.75 | 3.89 | -74.63 | -85.68 | -0.11 | -0.34 | -78.89 | -88.28 | -3.47 | -1.10 |
| 500 | TSSEM | -71.09 | -85.91 | 15.43 | 28.20 | -79.36 | -89.93 | 15.60 | 25.79 | -83.22 | -91.93 | 5.52 | 13.62 |
| | GLSr | -78.91 | -87.92 | -5.93 | 0.45 | -85.16 | -91.56 | -4.13 | -3.67 | -87.61 | -93.05 | -1.53 | -5.98 |
| | GLSm | -78.74 | -87.99 | -4.95 | 1.70 | -84.81 | -91.47 | -0.97 | -0.71 | -87.54 | -93.05 | -0.12 | -1.44 |
| 1000 | TSSEM | -79.36 | -89.93 | 15.82 | 25.93 | -85.46 | -92.73 | 10.20 | 22.82 | -87.75 | -94.12 | 16.64 | 17.05 |
| | GLSr | -84.98 | -91.51 | -3.07 | -3.29 | -89.26 | -93.97 | -4.29 | -5.01 | -91.21 | -95.08 | -0.70 | -3.95 |
| | GLSm | -85.00 | -91.56 | -1.47 | -1.35 | -89.27 | -93.99 | -3.74 | -3.18 | -91.20 | -95.08 | 0.19 | -4.27 |

*Note:* GLSr represents the original GLS method which uses sample correlations as substitute for population correlations in the asymptotic covariance matrix; GLSm represents the modified estimator of the asymptotic covariance matrix using mean correlation coefficient
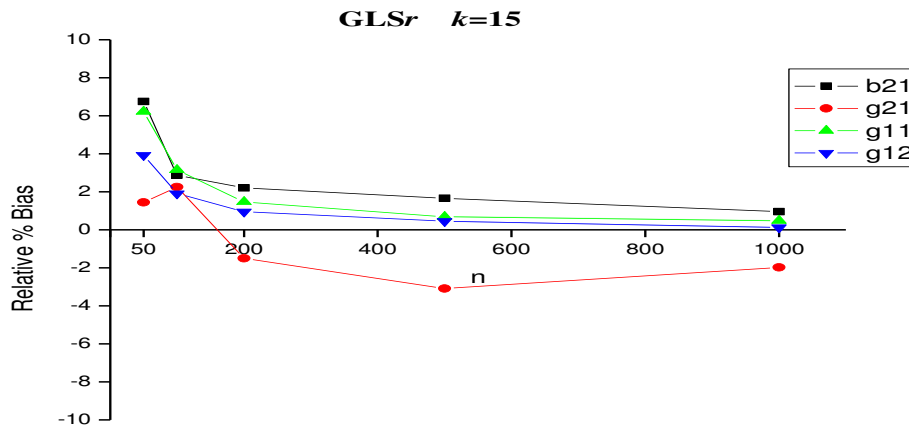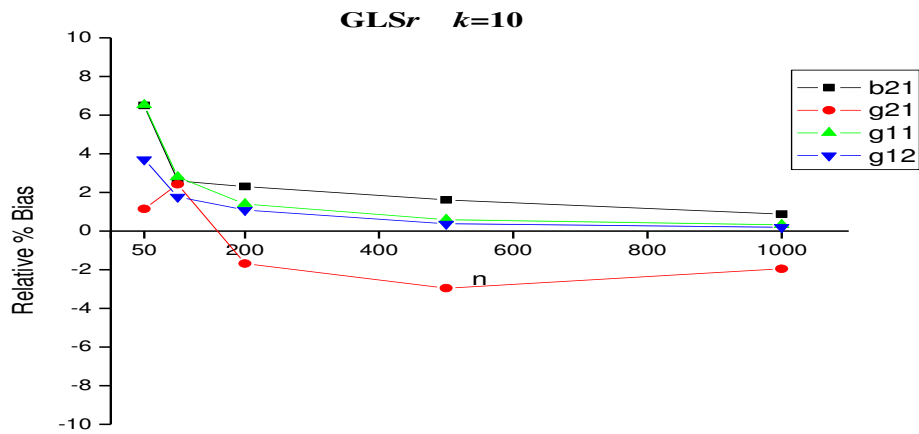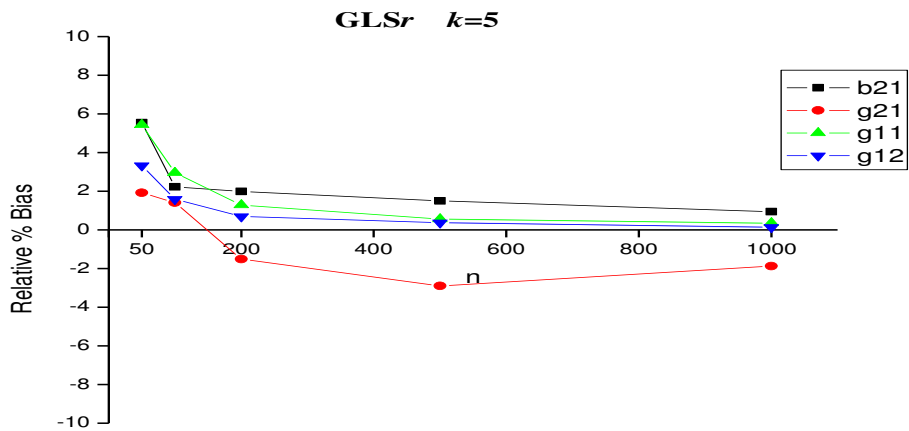
*Figure 4.11*. Relative % bias of the path coefficients for the GLS*r* approach in Study 2, fixed-effects analyses.
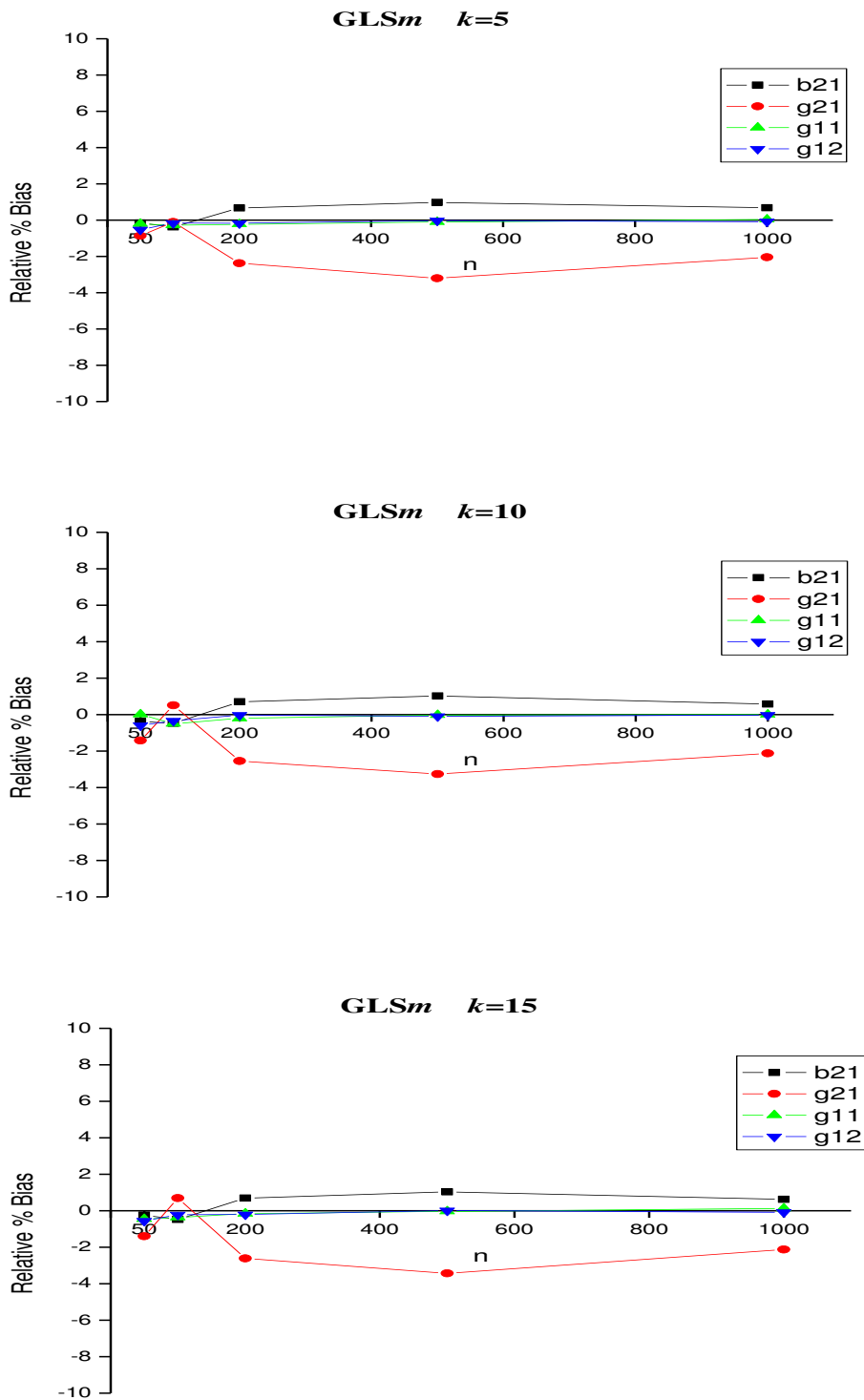
*Fure 4.12*. Relative % bias of the path coefficients for the GLS*m* approach in Study 2, fixed-effects analyses.
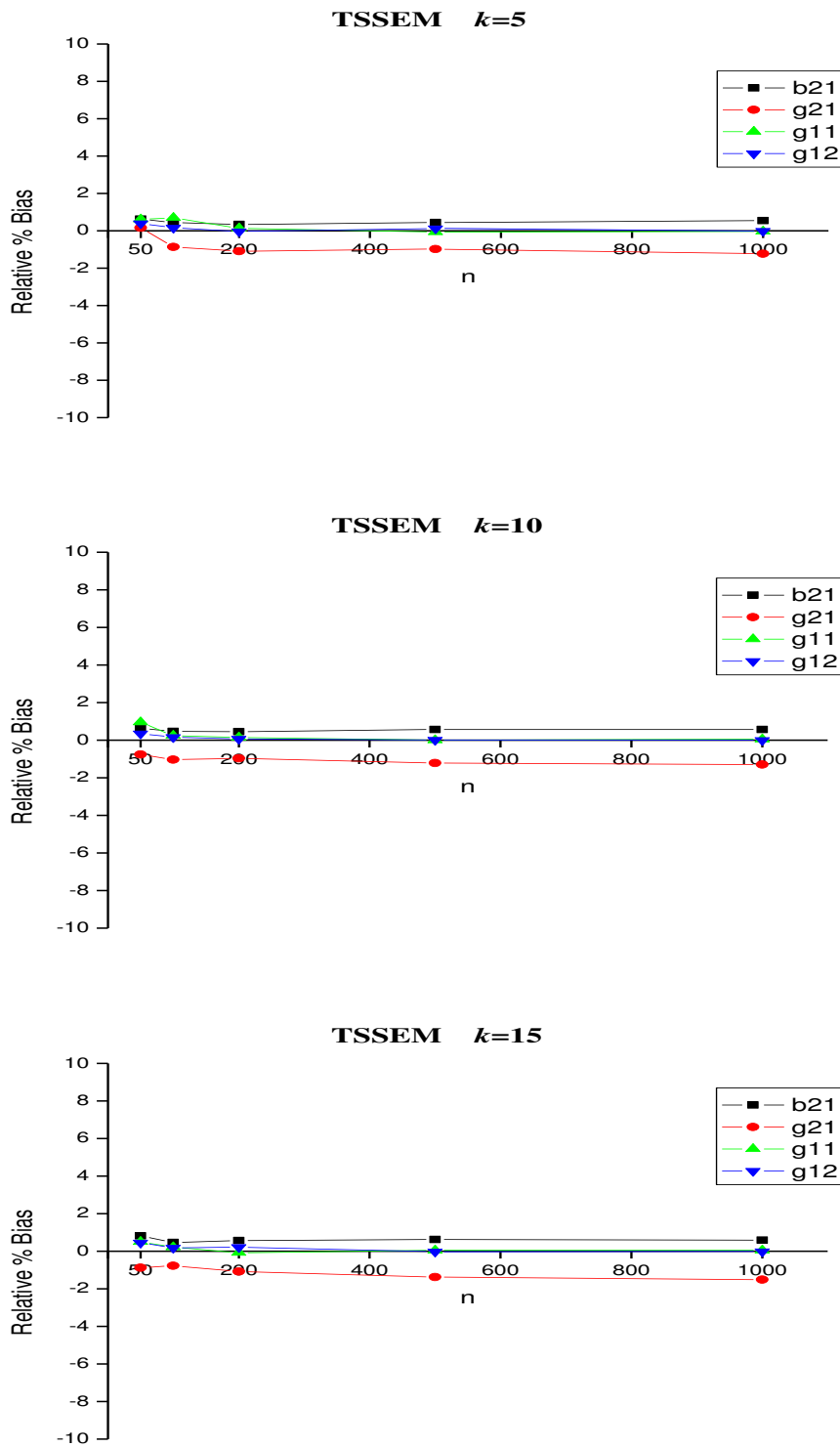
*Figure 4.13*. Relative % bias of the path coefficients for the TSSEM approach in Study 2, fixed-effects analyses.
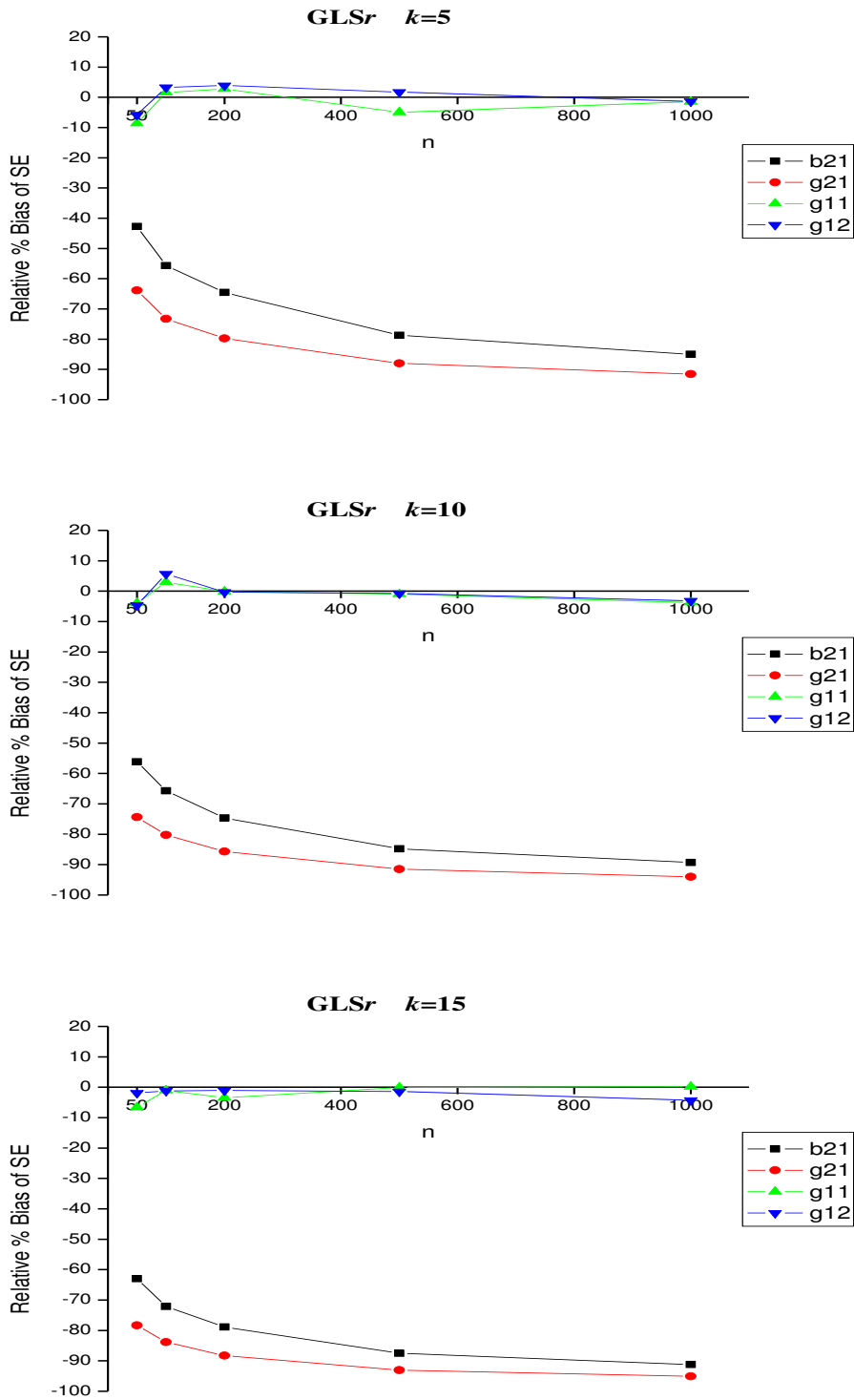
*Figure 4.14*. Relative % bias of the standard error of the path coefficients for the GLS*r* approach in Study 2, fixed-effects analyses.
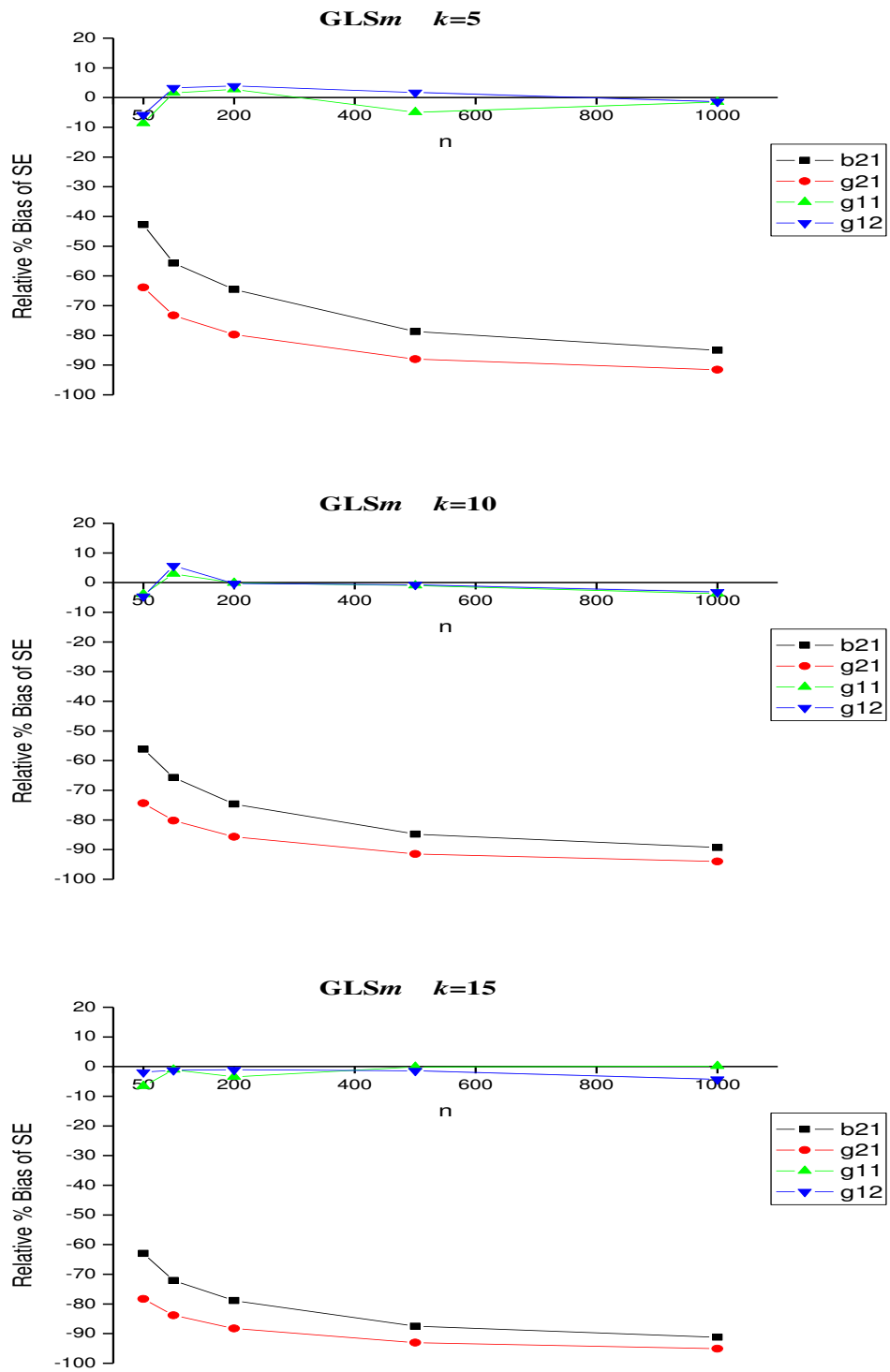
*Figure 4.15*. Relative % bias of the standard error of the path coefficients for the GLS*m* approach in Study 2, fixed-effects analyses.

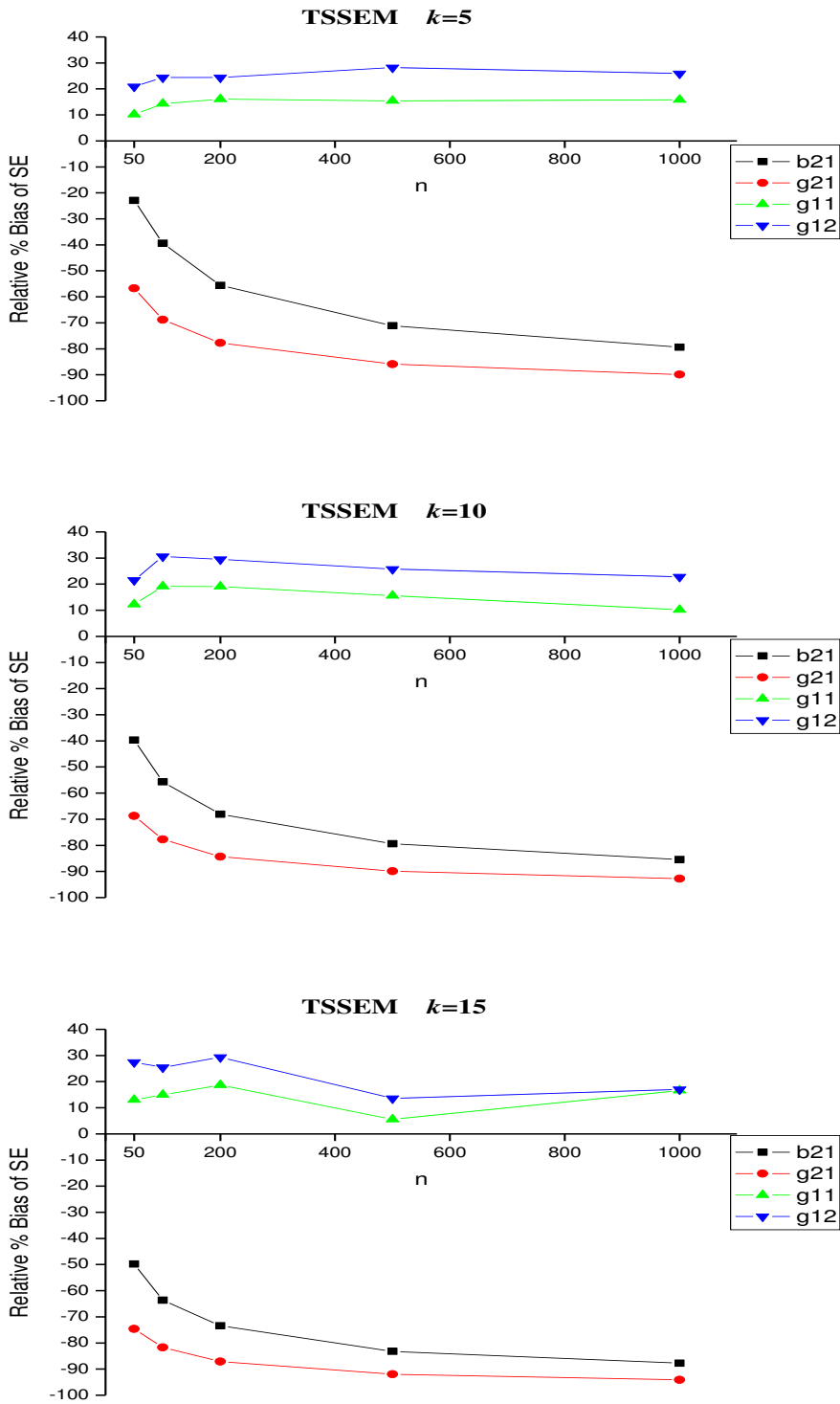*Figure 4.16*. Relative % bias of the standard error of the path coefficients for the TSSEM approach in Study 2, fixed-effects analyses.
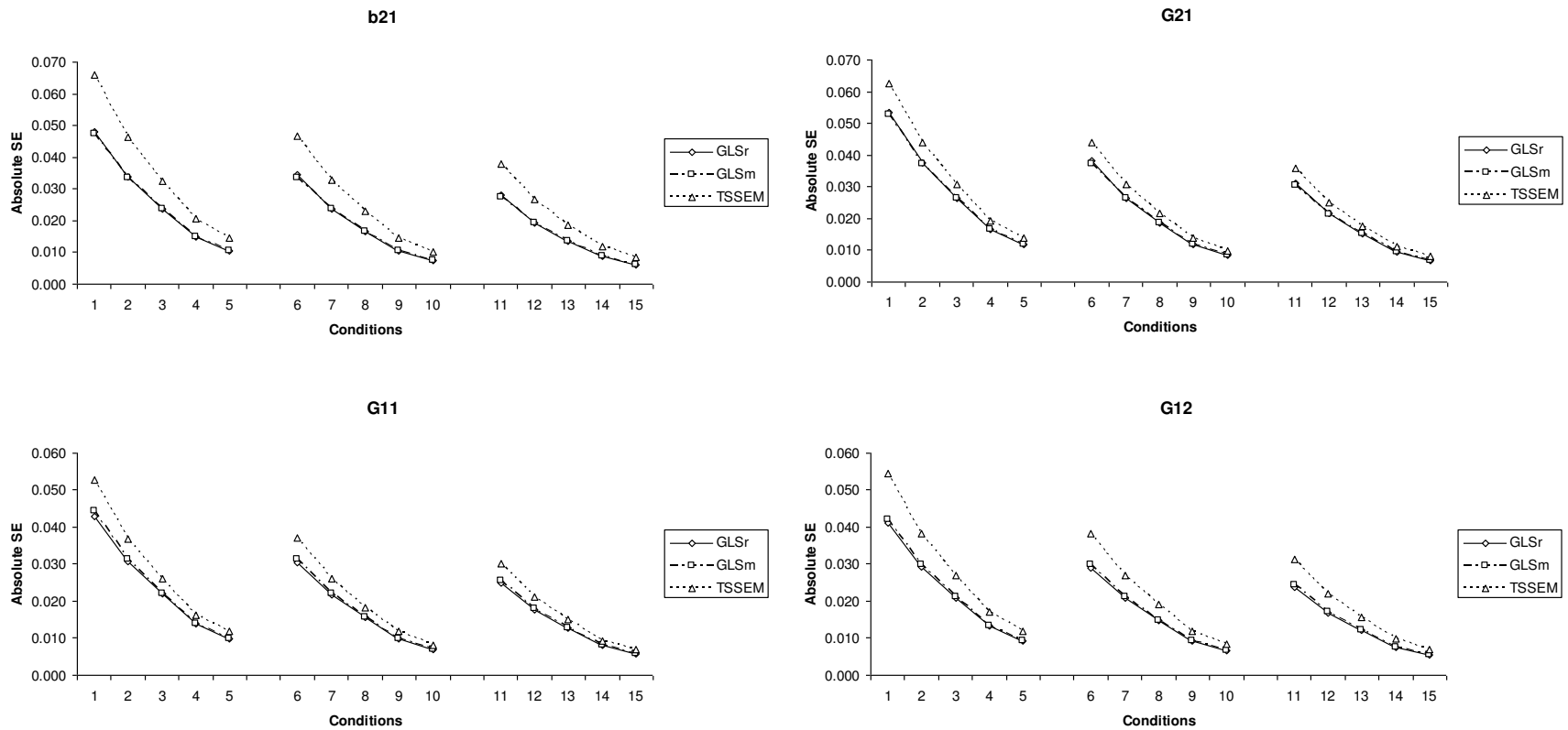
*Figure 4.17.* Mean standard errors of the path coefficients of each method in Study 2

**GLSm  k = 5**



**GLSm  k = 10**



**GLSm  k = 15**



*Figure 4.18*. Relative % bias of the path coefficient estimates for the GLS*m* approach in Study 2, random-effects analyses.

**GLSm k = 5**



**GLSm k = 10**



**GLSm k = 15**



*Figure 4.19*. Relative % bias of the standard error of the path coefficient estimates for the GLS*m* approach, random-effects analyses.
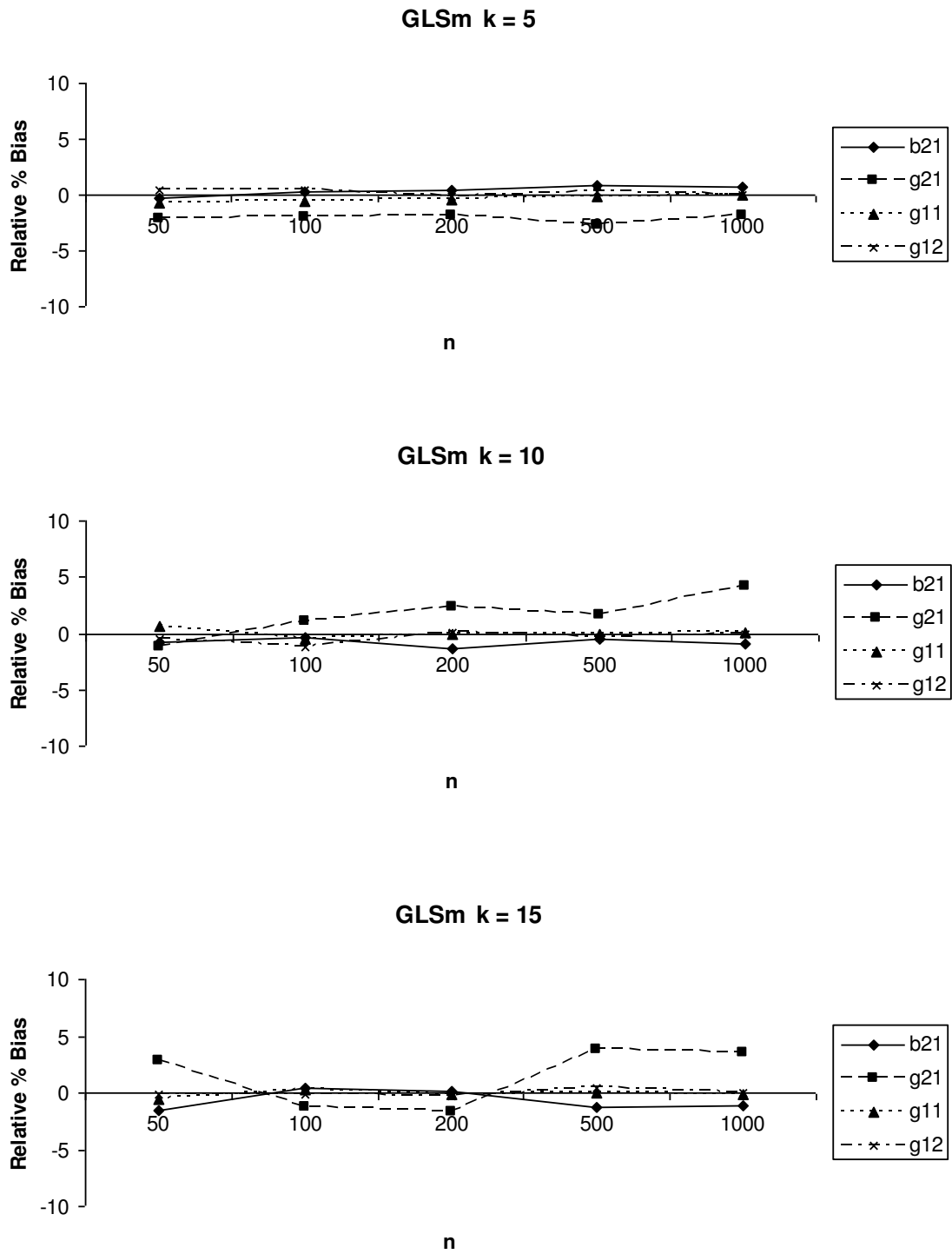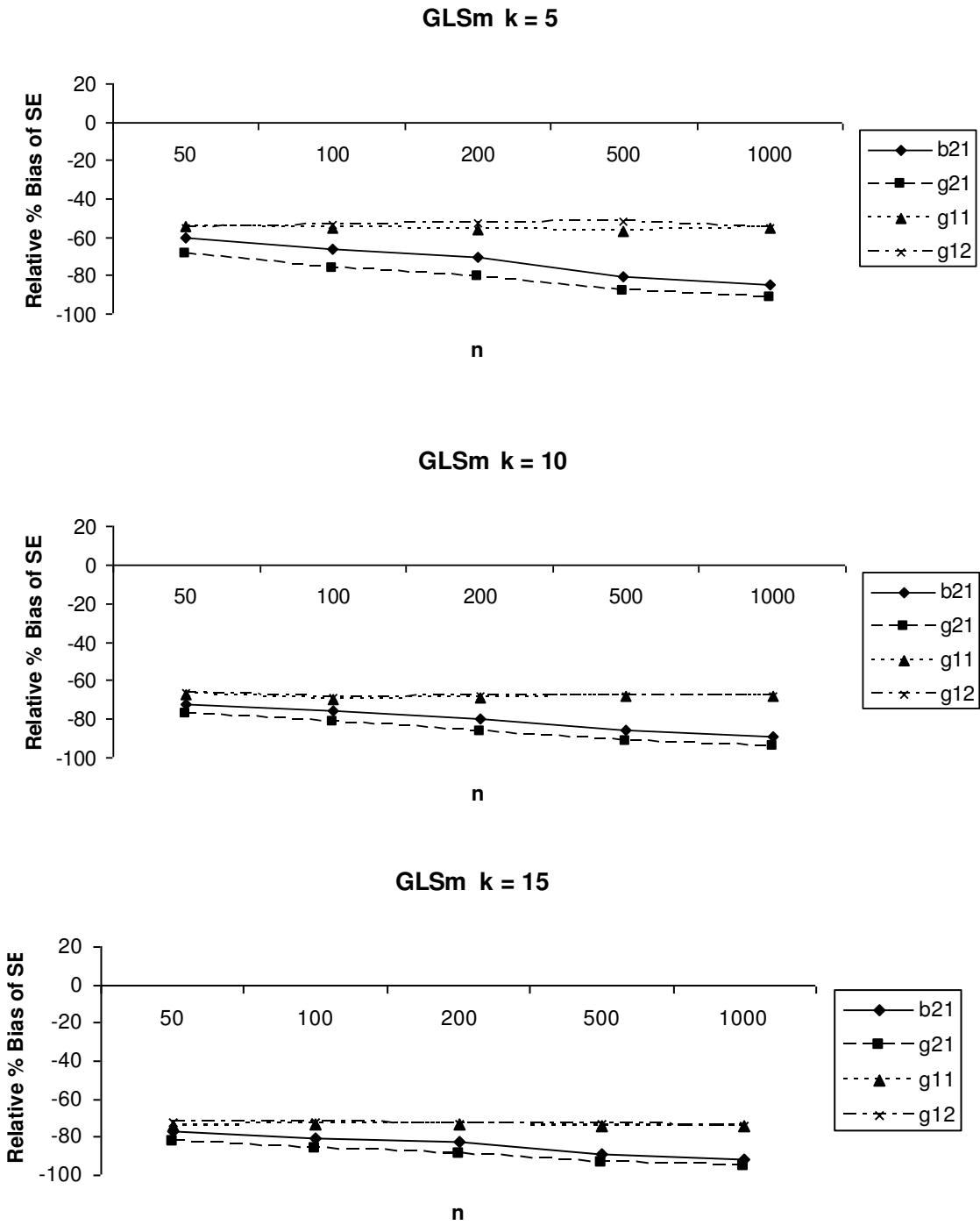
Table 5.1

*Sample Size and Correlation Coefficients*

| Sample ID | Sample Size $N$ | Cog_Per $r_{21}$ | Som_Per $r_{31}$ | Self_Per $r_{41}$ | Cog_Som $r_{32}$ | Cog_Self $r_{42}$ | Som_Self $r_{43}$ |
|---|---|---|---|---|---|---|---|
| 1 | 51 | -0.52 | -0.43 | 0.16 | 0.57 | -0.18 | -0.26 |
| 2 | 51 | 0.16 | -0.47 | 0.18 | -0.59 | -0.37 | -0.22 |
| 3 | 24 | -0.28 | -0.53 | 0.39 | 0.67 | -0.36 | -0.72 |
| 4 | 24 | -0.28 | -0.49 | 0.51 | 0.67 | -0.41 | -0.72 |
| 5 | 24 | 0.21 | -0.27 | 0.52 | 0.13 | -0.04 | -0.59 |
| 6 | 24 | -0.15 | -0.4 | 0.35 | 0.67 | -0.36 | -0.72 |
| 7 | 24 | -0.24 | -0.24 | 0.36 | 0.67 | -0.41 | -0.72 |
| 8 | 24 | -0.06 | -0.16 | 0.22 | 0.13 | -0.04 | -0.59 |
| 9 | 100 | 0.23 | 0.08 | 0.51 | 0.45 | -0.29 | -0.44 |
| 10 | 100 | 0.08 | 0.08 | 0.35 | 0.45 | -0.29 | -0.44 |
| 11 | 199 | 0.58 | 0.57 | 0.55 | 0.6 | -0.57 | -0.68 |
| 12 | 128 | 0.14 | 0.02 | 0.13 | 0.56 | -0.53 | -0.27 |
| 13 | 14 | -0.39 | -0.17 | 0.19 | 0.21 | -0.54 | -0.43 |
| 14 | 142 | -0.55 | -0.48 | 0.66 | 0.47 | -0.38 | -0.46 |
| 15 | 37 | 0.53 | -0.12 | 0.03 | 0.52 | -0.48 | -0.40 |
| 16 | 58 | -0.29 | -0.36 | 0.41 | 0.55 | -0.6 | -0.69 |
| 17 | 70 | -0.09 | -0.29 | 0.39 | 0.62 | -0.46 | -0.54 |
| 18 | 70 | -0.01 | -0.16 | 0.42 | 0.62 | -0.46 | -0.54 |
| 19 | 30 | -0.27 | -0.13 | 0.15 | 0.63 | -0.68 | -0.71 |
| 20 | 56 | -0.76 | -0.63 | 0.63 | 0.76 | -0.65 | -0.61 |

*Note.* Some of the correlation vectors are from the same primary study. In this dissertation, they are assumed to be from independent samples. Cog-, Per-, Som-, and Self- represent cognitive anxiety, performance, somatic anxiety, and self-confidence, respectively. Thus, Cog_Per refers to the correlation between cognitive anxiety and performance; Som_Per represents the correlation between somatic anxiety and performance; Self-Per is the correlation between self-confidence and performance; Cog-Som is the correlation between cognitive anxiety and somatic anxiety; Cog-Self is the correlation between cognitive anxiety and self-confidence; Som-Self is the correlation between somatic anxiety and self-confidence.

Table 5.2

*Pooled Correlation Matrix by the GLSr, GLSm and TSSEM approach*

*-- Fixed-effects Analysis*

| Method | Cog-Per | Som-Per | Self-Per | Cog-Som | Cog-Self | Som-Self |
|---|---|---|---|---|---|---|
| GLSr | -.08 | -.09 | .47 | .53 | -.48 | -.68 |
| GLSm | -.01 | -.10 | .40 | .50 | -.43 | -.50 |
| TSSEM | -.06 | -.16 | .37 | .48 | -.40 | -.47 |

Table 5.3

*Path Coefficients and Standard Errors Obtained by the GLSr, GLSm and*

*TSSEM approach – Fixed-effects Analysis*

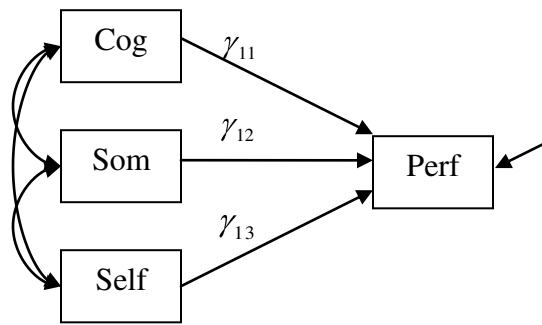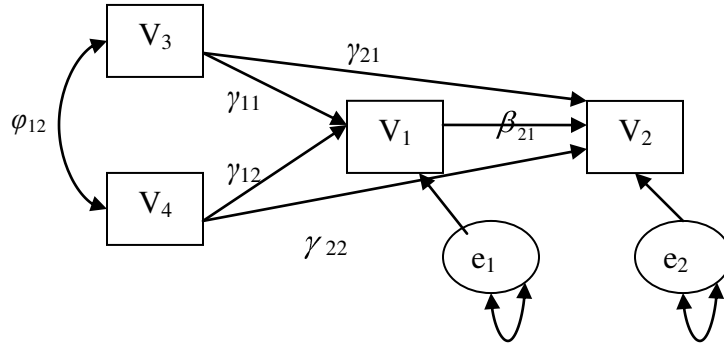| Method | $\hat{\gamma}_{11}$ ($SE_{11}$) | $\hat{\gamma}_{12}$ ($SE_{12}$) | $\hat{\gamma}_{13}$ ($SE_{13}$) |
|---|---|---|---|
| TSSEM | .108 (.031) | -.019 (.033) | .408 (.033) |
| GLS*m* | .180 (.030) | .069 (.031) | .516 (.028) |
| GLS*r* | .110 (.036) | .207 (.046) | .652 (.035) |

*Figure 5.1.* Path model showing the CSAI-2 and performance relationship

# APPENDIX A

**Direct Decomposition of the Correlation Coefficients in the Path Model**



$$r_{13} = \gamma_{11} + \gamma_{12} \times r_{12}$$

$$r_{14} = \gamma_{12} + \gamma_{11} \times r_{12}$$

$$r_{23} = \gamma_{21} + \beta_{21} \times \gamma_{11} + \beta_{21} \times \gamma_{12} \times r_{12}$$

$$r_{24} = \gamma_{22} + \beta_{21} \times \gamma_{12} + \beta_{21} \times \gamma_{11} \times r_{12}$$

$$r_{12} = \beta_{21} + \gamma_{21} \times \gamma_{11} + \gamma_{22} \times \gamma_{12} + \gamma_{21} \times \gamma_{12} \times r_{12} + \gamma_{22} \times \gamma_{11} \times r_{12}$$

*Note.* The correlation coefficient $r_{23}$ was treated as random in the random-effects data generation. As can be seen here that the most affected path coefficients are $\beta_{21}$ and $\gamma_{12}$.

# REFERENCES

Becker, B. J. (1992). Using results from replicated studies to estimate linear models. *Journal of Educational Statistics*, 17, 341-362.

Becker, B. J. (1995). Corrections to "Using results from replicated studies to estimate linear models." *Journal of Educational and Behavioral Statistics*, 20, 100-102.

Becker, B. J., & Schram, C. M. (1994). Examining explanatory models through research synthesis. In H. Cooper and L. V. Hedges (Eds). *The handbook of research synthesis* (pp. 357-381). New York: Russell Sage Foundation.

Becker, B. J., & Fahrbach, K. (1994, April). *A comparison of approaches to the synthesis of correlation matrices*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499-525). San Diego, CA: Academic Press.

Becker, G. (1996). The meta-analysis of factor analyses: An illustration based on the cumulation of correlation matrices. *Psychological Methods*, 1, 341–353.

Beretvas, S. N., & Furlow, C. F. (2006). Evaluation of an approximate method for synthesizing covariance matrices for use in meta-analytic SEM. *Structural Equation Modeling, 13*, 153-185.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. *Journal of Vocational Behavior, 72*, 298-308.

Browne, M.W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*, 62-83.

Cheung, M. W. L. (2007a). TSSEM: A LISREL syntax generator for two-stage structural equation modeling (Version 1.10) [Computer software]. Retrieved November 30, 2008, from http://courses.nus.edu.sg/course/psycwlm/internet/tssem.zip

Cheung, M. W. L., & Chan, W. (2004). Testing dependent correlation coefficients via structural equation modeling. *Organizational Research Methods*, 7, 206–223.

Cheung, M. W. L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. Psychological Methods, 10, 40–64.

Cheung, M. W. L. (2007). TSSEM: A LISREL syntax generator for two-stage structural equation modeling (Version 1.10) [Computer software and manual]. Retrieved   from http://courses.nus.edu.sg/course/psycwlm/internet/tssem.zip.

Cheung, S. F. (2000). Examining solutions to two practical issues in meta-analysis: Dependent correlations and missing data in correlation matrices. Dissertation Abstracts Internaltional, 61, 8B. (UMI NO. AA199884691)

Cooley, W. W., & Lohnes, P. R. (1971). *Multivariate data analysis*. New York: Wiley.

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology, 85*, 678–707.

Colquitt, J. A., Scott, B. A., & LePine, J. A. (2007). Trust, trusworthiness, and trust Propensity: a meta-analytic test of their unique relationships with risk taking and job performance. *Journal of Applied Psychology, 92* (4), 909-927.

Cooper, H. M., & Hedges, L. V. (Eds.). (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

Craft, L. L., Magyar, T. M., Becker, J. B., & Feltz, L. D. (2003). The relationship between the Competitive State anxiety Inventory-2 and sport performance: a meta-analysis. *Journal of Sport & Exercise Psychology, 25*, 44-65.

Cudeck, R. (1989). Analysis of correlation matrices using covariance structure models. *Psychological Bulletin, 105*, 317-327.

Field, A. P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects models. *Psychological Methods, 6,* 161-180.

Field, A. P. (2005). Is the meta-analysis of correlation coefficients accurate when population correlations vary? *Psychological Methods, 10,* 444-467.

Fisher, R.A. (1921). "On the 'probable error' of a coefficient of correlation deduced from a small sample". *Metron* 1: 3–32.

Fouladi, R. T. (2000). Performance of modified test statistics in covariance and correlation structure analysis under conditions of multivariate nonnormality. *Structural Equation Modeling, 7*, 356-410.

Furlow, C. F., & Beretvas, S. N. (2005). Meta-analytic methods of pooling correlation matrices for structural equation modeling under different patterns of missing data. *Psychological Methods, 10*, 227–254.

Gagne, P, & Furlow, C. F. (2009). Automating multiple software packages in simulation research for structural equation modeling and hierarchical linear modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16 (1),* 179-185.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*, 3–8.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. M. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis (pp. 339 – 335). New York: Russell Sage.

Hafdahl, A. R. (2001). *Multivariate meta-analysis for exploratory factor analytic research.* Unpublished doctoral dissertation, University of North Carolina, Chapel Hill.

Hafdahl, A. R. (2007). Combining correlation matrices: Simulation analysis of improved fixed-effects methods. *Journal of Educational and Behavioral Statistics, 32*, 180-205.

Hafdahl, A. R., & Williams, M. A. (2009). Meta-analysis of correlations revisited: Attempted replication and extension of Field's (2001) simulation studies. *Psychological Methods, 14(1),* 24-42.

Hafdahl, A. R. (2009). Improved Fisher z estimators for univariate random-effects meta-analysis of correlations. *British Journal of Mathematical and Statistical Psychology, 62*, 233-251.

Hedges, L. V., & Olkin, I. (1985). Statistical methods for meta-analysis. Orlando, FL: Academic Press.

Hedges, L. V, & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3,* 486-504.

Hoogland, J. J., & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and meta-analysis. *Sociological Methods and Research,* 26, 329-367.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.

Jennrich, R. I. (1970). An asymptotic $\chi^2$ test for the equality of correlation matrices. *Journal of the American Statistical Association, 65,* 904 – 912.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: User's guide.* Chicago: Scientific Software.

Jöreskog, K. G., & Sörbom, D. (2006). *LISREL 8.8* [Computer software]. Chicago, IL: Scientific Software International, Inc.

Jöreskog, K. G., Sörbom, D., Du Toit, S., & Du Toit, M. (1999). *LISREL 8: New statistical features.* Chicago: Scientific Software International.

Kaplan, D. (2000). *Structural equation modeling: foundations and extensions*. Thousand Oaks, CA: Sage.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York: Guilford.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Martens, R., Vealey, R. S., & Burton, D. (1990a). *Competitive anxiety in sport*. Chanmpaign IL: Human Kinetics.

Martens, R., Burton, D., Vealey, R. S., Bump, L. A., & Simith, D. E. (1990b). Development and validation of the Competitive State Anxiety Inventory-2 (CSAI-2). In R. Martens, R. S. Vealey, & D. Burton (Eds), *Competitive anxiety in sport* (pp. 193-208). Champaign, IL: Human Kinetics.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus user's guide* (3rd ed.). Los Angeles, CA: Muthén & Muthén.

National Research Council. (1992). *Combining information: Statistical issues and opportunities for research.* Washington, DC: National Academy Press.

Olkin, I., & Siotani, M. (1976). Asymptotic distribution of functions of a correlation matrix. In S. Ideka (Ed.), *Essays in probability and statistics* (pp. 235-251). Tokyo: Shinko Tsusho.

Pigott, T. (1994). Methods for handling missing data in research synthesis. In H. Cooper & L. V. Hedges (Eds.), The handbook of research synthesis (pp. 163-175). New York: Russell Sage Foundation.

Raudenbush, S. W., Becker, B. J., & Kalaian, H. (1988). Modeling multivariate effect sizes. *Psychological Bulletin, 103*, 111−120.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Raykov, T. (2001). Testing multivariate covariance structure and means hypotheses via structural equation modeling. *Structural Equation Modeling*, *8*, 224-256.

SAS Institute. (1999). *SAS procedures guide, version 8*. Cary, NC: Author.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71*, 432–439.

Shadish, W. R. (1996). Meta-analysis and the exploration of causal mediating processes: A primer of examples, methods, and issues. *Psychological Methods, 1*, 47–65.

Specht, D. A. (1975). On the evaluation of causal models. *Social Science Research,4,* 113-133.

Silver, N. C., & Dunlap, W. P. (1987). Averaging correlation coefficients: Should Fisher's Z transformation be used? *Journal of Applied Psychology, 72,* 146-148.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245-251.

Sutton, A.J., Abrams, K. R., Jones, D. R., Sheldon, T. A., Song, F. (2000). *Methods for meta-analysis in medical research*. London: John Wiley.

Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometricmeta-analysis and structural equations modeling. *Personnel Psychology, 48,* 865–885.

# BIOGRAPHICAL SKETCH

Ying Zhang is from Anhui province of China. She received a Bachelor degree of Science in Geography Education from Anhui Normal University in 1997. She received her first Master of Science degree in Physical/Biological Geography from Anhui Normal University in 2000. In 2004, she received her second Master of Science degree in Measurement and Statistics from Florida State University. During the graduate study at FSU, she has been working with Dr. Betsy Becker, Dr. Yanyun Yang and Dr. Akihito Kamta on projects of meta-analysis, structural equation modeling, reliability and other measurement issues. She has been a teaching assistant for different graduate courses in the Measurement and Statistics program, such as ANOVA, GLM, Multivariate Data Analysis, SEM etc. During the last year of her pursuing of her doctorate degree, she works for NCS Pearson Inc. as a statistical analyst on the Florida Comprehensive Test (FCAT) project with the Office of Assessment and School Performance at the Florida Department of Education. Her research interests are meta-analysis, structural equation modeling, reliability theory and large-scale assessments.