2013-01-01

Multiplicity Adjustments For Respecification Searches In Structural Equation Models

John Appiah Kubi
University of Texas at El Paso, luckubi@yahoo.com

MULTIPLICITY ADJUSTMENTS FOR RESPECIFICATION SEARCHES IN STRUCTURAL EQUATION MODELS

JOHN APPIAH KUBI

Department of Mathematical Sciences

APPROVED:

Amy E. Wagler, Chair, Ph.D.

Joan G. Staniswalis, Ph.D.

Thompson Sarkodie-Gyan, Ph.D.

Benjamin C. Flores, Ph.D.
Dean of the Graduate School

to my

MOTHER and FATHER

with love

MULTIPLICITY ADJUSTMENTS FOR RESPECIFICATION SEARCHES IN STRUCTURAL EQUATION MODELS

by

JOHN APPIAH KUBI

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Mathematical Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

August 2013

Acknowledgements

I would like to express my greatest gratitude to my advisor, Dr. Amy Wagler , for her advice, encouragement, enduring patience and indefatigable support throughout this task. I also wish to thank the other members of my committee, Dr. Staniswalis and Dr. Sarkodie-Gyan profusely for their valuable comments, suggestions and extreme patience in the face of numerous impediments. A special thanks of mine also goes to all the Professors and staff at the University of Texas at El Paso Mathematical Sciences Department for their dedication , support and guidance which has helped me to complete my degree and shaped me into who I am today.

Last but not the least, I want to thank my family and friends who appreciated me for my work and motivated me and finally to God Almighty who made all things possible

Abstract

Structural Equation Modeling (SEM), as a statistical modeling technique, is one of the most comprehensive and flexible approaches to data analysis currently available. Its use has been increasing steadily over the past few decades. Generally, it refers to a family of techniques that employs the analysis of covariance to establish relationships among a set of variables. It allows researchers (or users) to assess the adequacy of their hypothesized models with their sample data. Often times, in assesing their models, researchers are not only interested in the overall fit of their model but they are also interested in knowing which proposed relationships (parameters) are significant. With respect to the evaluation of the significance of parameters, researchers risk capitalizing on chance and including unnecessary parameters (that is, producing a less parsimonious model) in their models when no form of controlling type I error rate is adopted.

In this thesis, the Functional Independence Measure (FIM) data was used to demonstrate the effectiveness of using a Scheffe-like procedure for controlling the rate of type I errors when multiple parameters are evaluated for significance.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Structural Equation Modeling (SEM), as a statistical modeling technique, has been widely used over the past few decades. It is a technique that explores complex relations between one or more independent variables and one or more dependent variables. SEM encompasses a broad array of models, from linear regression to measurement models to simultaneous equations, including confirmatory factory analytic (CFA) models, latent growth models, etc.

A structural equation model is initially specified in accordance with a theoretical conception of the relationship between the observed (measured) variables and unobserved (latent) variables. Primarily, SEM is used to assess how a hypothesized model fits their data adequately. Users of SEM often times discover that their hypothesized model does not fit the data perfectly and so conduct specification searches for better fitting models (Joreskog, 1993). The searches may involve deletion and/ or addition of parameters that improves model fit. With respect to which parameters to add or delete, two alternatives have dominated the literature.

The first option utilizes residual statistics and sequential model modification indices (eg. Wald Test, Lagrange Multiplier Test) to identify which parameters to delete from or add to the model to improve the fit of the model to the sample data. This option is exploratory and is used when the goal of the analysis is to derive or improve a theoretical model for future evaluation. The second option is confirmatory (assuming a satisfactory fit of the model) and evaluates the significance of each hypothesized parameter at a specified significance level

(α), declaring statistically significant parameters important and non-significant parameters unimportant, within the framework of the model.

A common problem with both of the aforementioned methods is that they risk capitalizing on chance when evaluating the significance of multiple parameters in a structural model. Several methods have been proposed in solving this problem. In this thesis, we seek to use a Scheffe-like method to reduce the problem of capitalization on chance in respecifying our model and also make post-hoc testing in SEM more acceptable and consistent theoretically.

1.2 Significance of the study

SEM is more of a confirmatory modeling technique than exploratory in the sense that, it is mainly used to assess the adequacy of a predetermined theoretical model and to explain relationships among observed (manifest) variables and unobserved (latent) variables. Unfortunately, the theoretical model does not always fit the data perfectly and substantive changes or revisions to the model are necessary. This process is called respecification. In the outset, the number of respecifications are unknown, but a very large number are possible. A Lagrange Multiplier Statistic (Modification Index) is used to test whether a particular respecification improves the fit. Given that a large number of these MI tests may be performed, a method for controlling the overall Type I error is warranted. This study explores the use of a Scheffe-like adjustment for making respecification to SEMs.

1.3 Research questions

This research seeks to cover the following areas in its analysis:

- What is the structure and the dimensionality of the FIM Model ?
- How do we control multiplicity in relation to the number of model modifications and the criteria adopted for declaring some parameters significant ?

1.4 FIM

The FIM is the most widely accepted functional assessment measure in use in the rehabilitation community. It is an 18-item, 7-level ordinal scale with 1 the lowest possible rating indicating total dependence and 7 the highest possible rating indicating total independence. The tool is designed to measure " burden of care ", or " the type and amount of assistance" required for a person with a disability to perform basic life activities (Deutsch, Brawn and Granger, 1996 p.268). A FIM score of 2 means that the person puts forth less than 50% of effort necessary to do a task, but at least 25%. A score of 3 means " moderate assistance ", in which the person puts forth between 50% and 74% of the effort necessary to perform a task, and requires no more than helping or touching. A score of 4 means " minimal contact assistance ", in which the person puts forth 75% or more of the effort necessary to do a task, and requires no more help than touching.

A score of 5 means " supervision or setup " , in which the person only needs someone to standby and cue or coax him/her (without physical contact) so that he/she can do a task. A score of 5 can also be obtained if a helper is needed to set up items or assistive devices for the person. If someone gets a score of 3, 4 or 5 on the FIM scale, he/she is classified as having " modified dependence " , because the person can at least put forth half or more of the energy to complete the task.

A score of 6 on the FIM scale means " modified independence " , in which no helper is needed and the person needs an assistive device. A score of 6 can also be obtained when no help is needed but the person takes more than a reasonable amount of time to do a task or may complete the task in an unsafe manner. If someone gets a score of 6 or 7 on the FIM scale, he/she is classified as being Independent, because another person is not needed to complete the activity.

The 18-items of the FIM are composed of 13 motor tasks and 5 cognitive tasks. By adding the points for each item, the scores range from 18 (lowest) to 126 (highest) level of function. Scores are generally rated at admission and discharge. The motor dimension in-

cludes items assessing eating, grooming, bathing, upper body dressing, lower body dressing, toileting, bladder management, bowel management, bed to chair transfer, toilet transfer, shower transfer, locomotion (ambulatory or wheelchair level), stairs activities. The cognitive dimension consists of items assessing comprehension, expression, social interaction, problem solving and memory. These FIM areas are considered activities of daily lives which are activities one performs in the course of daily life.

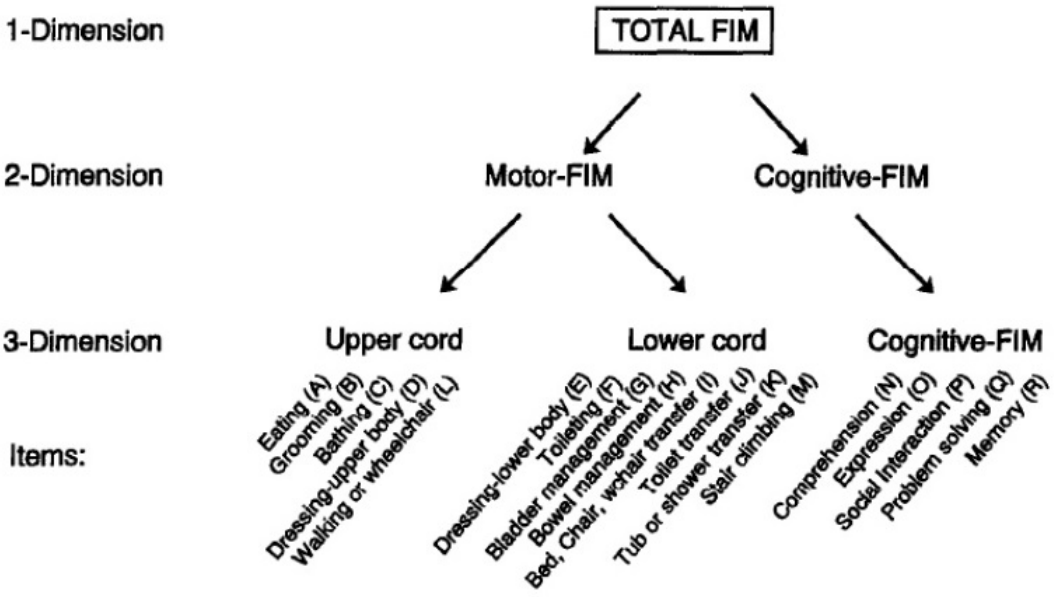

Figure 1.1: Dimensions within the FIM, Stineman et al., (1997)

1.5 How FIM scores are used

Patients recovering from cerebrovascular diseases (CVDs) or strokes are usually hospitalized for medical care and stabilization, investigations and rehabilitation if needed. For proper reintegration of these poststroke patients into society, it is vital to understand the difficulties or challenges they are faced with once they have been discharged from the hospital. The

difficulties they face vary from person to person and the ability to quantify the functional ability or status of these patients may be helpful in predicting their functional outcome. The FIM is widely used to measure the progress of functional skills or status of patients throughout the rehabilitation process. Although, limited evidence exists, the FIM scores are an accurate predictor of outcomes in poststroke or CVD patients. Since its introduction about 20 years ago, it has been known to serve as a consistent data collection tool for comparison of rehabilitation outcomes across the continuum of healthcare. It has been deemed to be a valid tool by clinicians and health experts because of the extent to which it predicts outcomes in medical rehabilitation. This form of validity is often known as predictive validity and it's representative of how well a scale predicts criterion scores. It is the best predictor of minutes of help needed per day for patients with multiple sclerosis, stroke and traumatic brain injury.

Chapter 2

Literature Review

2.1 General Overview-Latent variables

Latent variable modeling has gradually gained roots and prominence in mainstream statistics. Its applications extend to many disciplines and subject areas. Latent variables have been interpreted differently in different disciplines although their models have a similar mathematical structure. For example, latent variables are often random effects, common factors, latent classes, frailties, etc. They can simply be defined as random variables that are not directly observed or whose realizations are hidden but are rather inferred from other variables that are observed or measured directly.

Latent variable modeling, as indicated above is currently used in different areas including longitudinal data analysis, covariate measurement error, multivariate survival, market segmentation, psychometric measurements, meta-analysis etc.

As it is with any emerging field, latent variable modeling has had its own share of skepticism and prejudice. These cynics and critics see latent variable modeling as a dubious exercise filled with unverifiable assumptions and naive inferences regarding causality. These viewpoints have been nullified on at least three grounds:

1. That any reasonable statistical method can be abused by naive model specifications and over-enthusiastic interpretation ;

2. Ignoring latent variables often implies stronger assumptions than including them ;

3. Many of the assumptions in latent variable modeling can be assessed empirically and some can be relaxed if need be.

Latent variables are used to represent various phenomena such as "true" variables measured with error, hypothetical constructs unobserved heterogeneity, missing data, conterfactuals or potential outcomes, and latent responses underlying categorical variables

2.2 "True" variables measured with error

Latent variable can be used to represent a "true" variable which is measured with error. This concept is based on classical test theory (e.g. Lord and Novick, 1968) which assumes that any measure is a function of two variables: the true score and the error variation. This can be mathematically written as

$$y_j = \eta_j + \epsilon_j \tag{2.1}$$

where y_j represents the observed score on the measure, η_j is the person's true score and ϵ_j is the error variation for the jth subject. Examples of 'true' variables measured with error are: self-reported weight, physical activity measured by self-report, lung capacity measured by FEVI (Forced Expiratory Volume in 1 second) etc. The errors have zero means and so the expected value of the observed score gives the true score.

In structural equations modeling (SEM), latent variables are designed to represent true scores. Confirmatory factor analysis models are visually represented in the following way:

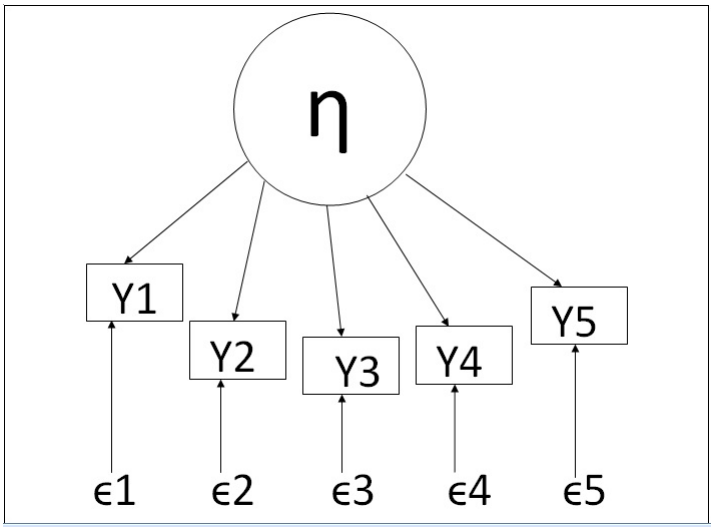

Figure 2.1: Path Diagram of the classical measurement model

.

The circle represents the latent variable, the rectangle represents the observed score (measurement) and the arrows represent linear relations. Measurement models are usually specified assuming continuous latent variables η_j. Such models are called factor models when the observed scores (measures) are continuous and item response models when the measures are categorical. Measurement models with both categorical latent and measured variables are known as latent class models.

The relationship between measurements (observed variables) and the latent variables (categorical or continuous) is direct and the measurements are assumed to be conditionally independent, that is, any covariance among the measurements is due to their common dependence on the latent variable and there should be no covariance between measurements along any other latent dimension.

Factor Analytic Models

In this section, we provide a brief overview of the common factor model and how it forms the basis for both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). We will also delve into the similarities and disparities between the EFA and CFA. The EFA and the CFA are used in conjunction with one another. The EFA is used in the early stages of scale development by exploring and defining the structure of a testing instrument while the CFA is used in the later stages of scale development. It is therefore imperative for an applied CFA researcher to have an idea as to how the EFA works.

The Common Factor Model

Factor analysis, since its introduction by Spearman (1987) has become a vital tool in multivariate statistical procedures and its application in research pervade across a host of disciplines (e.g., psychology, education, or sociology). The basic intent of this procedure is to determine the number and nature of latent variables (factors) that account for the variation and covariation among a set of observed measures, called *indicators*. A factor, specifically is an unobservable and immeasurable variable that influences more than one observed measure and is responsible for correlations among these observed measures, that is, these observed measures are correlated with each other due to the fact that they have a common factor. This implies that taking into account this latent factor will result in zero correlation among these observed measures. In applied research, factor analysis is mostly used in psychometric evaluations or assessment of multiple-item testing instruments (e.g questionnaires, cf, Floyd and Widaman, 1995). For example, a reseracher might develop 20 or more items that he or she believes are indicators of the construct of self-esteem. During the early stages of scale development, he or she might use factor analysis to determine which of the 20 items are reasonable indicators of self-esteem, that is, which of them have strong correlations among themselves and also the construct of self-esteem.

In this sense, factor analysis can be used to reduce a larger set of intercorrelated indicators to a smaller set of composite variables. These concepts come from the common factor model which postulates that each indicator in a group of observed measures is a linear function of one or more common factors and one unique factor, that is, the variance of each indicator is partitioned into two parts;

- common variance, which is shared with other indicators as a result of their association with the common latent factor
- unique variance, which is a combination of reliable variable specific to the indicator and random error variance(measurement error)

Two main types of analyses exist based on the common factor model. These are Exploratory and Confirmatory data analysis. Both aim to reproduce observed relationships among indicators with a smaller set of latent variables. They differ in the number and nature of a priori specifications and restrictions made on the factor model. With EFA, no prior specifications and restrictions are made as to the number of latent variables and the sort of relationships that exist between indicators and the common factors. Rather, it is the researcher who employs exploratory data analysis to determine the appropriate number of common factors and discover which measured variables are reasonable indicators of the various latent dimensions. The goodness or otherwise of the prespecified factor solution is seen in how well it reproduces the sample correlation (covariance) matrix of the observed variables. This means that the CFA, unlike the EFA, requires a strong empirical or conceptual foundation to guide the specification and evaluation of the factor model. Understandably, the EFA is employed in the early stages of scale development and construct validation, whereas the CFA is used in the later stages after the underlying structure has been identified by the EFA on empirical and theoretical evidence.

Exploratory Factor Analysis

As stated earlier, an applied researcher of CFA needs to have a working knowledge of the EFA.We therefore need to discuss the procedures of the EFA in general. The main objective of the EFA is to evaluate the dimensionality of a group of multiple indicators by finding out the least number of factors needed to interpret or explain the intercorrelations among the indicators. Whereas, the researcher will eventually specify the number of factors, EFA is exploratory in that, no a priori restrictions are placed on the pattern of relationships between the observed measures and the latent variables. This is a major departure from the CFA. In CFA, the researcher specifies the major aspects of the model in advance including the number of factors and nature of relationship between the factors and indicators.

After careful analysis, if the researcher comes to the conclusion that the EFA is the best analytic technique for the data at hand, then he must decide which indicators to bring on board in the analysis and determine if the size and nature of the sample suffices for the analysis. Other procedural aspects of the EFA include:

- The choice of a specific method to estimate the model
- Selection of the appropriate number of factors
- In the case of models that have more than one factor, selection of a technique to rotate the initial factor matrix to enhance interpretation of the solution
- If desired, choosing a method to evaluate factor scores

Factor extraction

Many methods come to mind when we want to estimate the common factor model. Some of these methods include maximum likelihood, principal factors, weighted least squares, unweighted least squares, generalized least squares, imaging analysis, minimum residual analysis, alpha factoring etc. If the EFA has continuous indicators, the most frequently

used methods are maximum likelihood (ML) and principal factors (PF). The ML is also the most commonly used in CFA. Using the ML for factor extraction is advantageous because it allows for a statistical evaluation of how well the factor solution reproduces the relationships among the indicators in the input data; that is, how closely do the correlations among the indicators predicted by the factor analysis parameters approximate the relationships present in the input correlation matrix? The ML method more appropriately determines the appropriate number of factors. However, in using the ML extraction method, we need to assume that the variables have a multivariate normal distribution. If the input data do significantly deviate from the multivariate normal distribution, then the results from the ML method can be distorted, misleading and untrustworthy. (e.g. goodness of model fit, significance tests of model parameters). When this happens, then the PF extraction method is helpful because it is free from any distributional assumptions. This means PF methods are preferred in scenarios where the input data depart markedly from normality. This method however does not provide goodness-of-fit indices useful in determining the suitability of the factor model.

Factor Selection

When the factor analysis is run using any extraction method (e.g. PF, ML), the initial results are used to determine the appropriate number of factors to be extracted in subsequent analysis. This stage is considered as the most crucial stage in EFA because "under factoring"(selecting too few factors) or "over factoring"(selecting too many factors) can severely affect the validity of the factor model and invariably affect the resulting estimates (e.g. introduce appreciable error in the factor loading estimates), though the consequences of over factoring is less severe than that of under factoring. The decision about the appropriate number of factors should be guided by substantive considerations and some statistical guidelines.

It is important to note that the number of factors m that can be extracted by EFA

is limited by the number of observed measures p that are submitted to the analysis. The upper limit on the number of factors is relative to the extraction method used. For instance, in PF EFA, the maximum number of parameters that can be extracted is defined to be p - 1. In ML EFA, the number of parameters that are estimated in the factor solution (a) must be equal to or less than the number of elements, and (b) in the input covariance matrix or correlation matrix (i.e $a \leq$b).

Introduction to the CFA

Here we introduce the concept of the CFA in terms of its parameters and fundamental equations. An overview of the common factor model and EFA have been discussed in previous chapters and so we discuss the CFA here and make comparisons thoroughly with the EFA.

Similarities and Differences of EFA and CFA

The CFA's purpose, like the EFA is to identify latent factors that account for the variation and covariation among a set of observed measures. Both EFA and CFA are based on the common factor model and so they have some concepts in common. (e.g. factor loadings, unique variances, residuals and commonalities). However, they are different with respect to how the latent factors are identified in relation to the observed measures. While the EFA is in general, a descriptive or exploratory procedure, the CFA prespecifies all aspects of the factor model, that is, it specifies the number of factors, the pattern of indicator-factor loadings and so on. As it was stated in the previous chapters, the CFA requires a very strong empirical or conceptual foundation to guide the researcher on how well to specify and evaluate the factor model. Accordingly, the CFA is typically applied in later stages of scale development or construct validation after the EFA has been used to identify the underlying structure by prior empirical analysis and theoretical backings.

Mostly, the EFA and CFA rely on the same estimation methods. When the ML estimator, which is a full information estimator, is used, the models which are produced by both the EFA and CFA are evaluated in terms of how well the solution reproduces the observed variances and covariances among the input indicators (that is, goodness-of-fit evaluation). Additionally, the quality of EFA and CFA models is partly determined by the size of the resulting parameter estimates (e.g. magnitude of factor loadings and factor intercorrelations) and how well each factor is represented by observed measures (e.g. number of indicators per factor, size of indicator commonalities, factor determinacy).

Standardized and Unstandardized Solutions

Traditionally, all variables used in the EFA are *completely standardized*, a result of conducting analysis in R. Both the latent factors and observed measures used are also completely standardized: factor variances equal 1.0; factor loadings are interpreted as correlations or standardized regression coefficients. Like the EFA, the CFA also produces a completely standardized solution but much of the analysis does not standardize the latent factors or the observed variables. Typically, the CFA analyzes a variance-covariance matrix (needed to produce an unstandardized CFA solution) or raw data that are used by the software program to produce an input variance-covariance matrix instead of using a correlation matrix (that is, a correlation matrix is a completely standardized variance-covariance matrix). This means the CFA input matrix is made up of indicator variances on the diagonal and indicator covariances in the off-diagonal. Apart from a completely standardized solution, the CFA's results also include an *unstandardized solution* (parameter estimates expressed in the original metrics of the indicators) and possibly a *standardized solution* (relationships involving unstandardized indicators and standardized latent variables) Many aspects of the CFA such as the standard errors and significance testing of model parameters are based on unstandardized estimates.

We can also include the unstandardized means of the indicators in CFA. This means,

contrary to the EFA, which deals with completely standardized values, the CFA may deal with analyzing both unstandardized variance-covariance structures and mean structures (as the result of standardization in EFA, indicator means are presumed to be zero) We can estimate the means of the latent factors and the intercepts of the indicators when indicator means are included as input in CFA.

The results of EFA is reported as completely standardized solutions and that of the CFA is also mostly reported as completely standardized solutions. SEM methodologists, however prefer reporting unstandardized solutions because the analysis is itself based on unstandardized variables and completely standardized variables may be misleading. For example, the true nature of variance and other relationships among the observed measures may be hidden when these variables have been standardized. When the original metric of variables is expressed in meaningful units, unstandardized estimates give a better interpretation and clearly convey the importance or substantive significance of effects.

2.3 Statistical Multiplicity

Multiple testing refers to the testing of more than one hypothesis at a time. It is a subfield of the broader field of multiple inference, or simultaneous inference, which includes multiple estimation as well as testing. The problem of multiplicity or multiple comparisons also occurs when one considers a set of statistical inferences simultaneously or infer on selected parameters only, where the selection depends on the observed values.

Multiplicity of data, hypotheses and analyses is a common problem in biomedical and epidemiological research. However, there seems to be a lack of knowledge about statistical procedures for multiple testing. In general, in testing any single hypothesis, conclusions based on statistical evidence are uncertain. If we perform one significance test at level α , the probability of rejecting the null hypothesis when it is true, thus committing a Type I error is the *comparisonwise error rate* α, also called *individual error rate*. This means that the probability of not rejecting the true null hypothesis is (1 - α). If k independent

tests are performed, the probability of not rejecting all k null hypotheses when in fact all are true is $(1-\alpha)^k$. Hence, the probability of rejecting at least one of the k independent null hypotheses when in fact all are true is the *experimentwise error rate* under the complete null hypotheses is EER $= 1 - (1-\alpha)^k$, called the *global level* or $familywise$ *error rate*. If the number k of tests increases, EER also increases. This means that when many hypotheses are tested, and each test has a specified type I error probability, the probability that at least some type I errors are committed increases, often sharply, with the number of hypotheses. This may have serious consequences if the set of conclusions must be evaluated as a whole.

In SEM settings, researchers are often exploring models with numerous parameters to be estimated and the probability that any parameter will be significant by chance increases as the number of parameters to be tested in the model increases. There is therefore the need for SEM users to adopt a strategy that controls the probability of falsely declaring parameters significant (that is, committing type I errors) when multiplicity exists, although there are a few arguments for NOT applying this method. According to Kaplan and Wenger (1993), there is often a high degree of intercorrelatedness between parameters in a model and so methods for controlling the type I error rate becomes overly conservative. Secondly, Kaplan and Wenger (1993) suggest that it is more important in tests of parameters to maximize power (control type II errors) because failure to include important parameters can bias parameter estimates. Lastly, another problem by Games (1971) and Hancocks Klockars (1996) is how to define a particular set (family) of hypothesis when controlling type I error rate in SEM.

First of all, the fact that parameters in a model are intercorrelated affects the conservativeness of the type I error control procedures. This can be seen in the size and pattern of the input covariance matrix. However, since researchers cannot tell the extent to which the parameters in their model are correlated, they must be conscious of the risk of falsely declaring parameters significant. So, apart from assuming that every parameter is perfectly correlated with the other, SEM researchers should be aware that independence between pa-

rameters increases the probability of making type I errors above the nominal significance level. Secondly, power is important in all statistical tests but it is not meaningful if the rate of type I errors are not controlled. This means it is important to maximize power because it ensures that relevant parameters are retained in the model but not at the expense of falsely retaining nonsignificant parameters. Lastly, a family of hypotheses is specified so that we do not select a set that is so large that it is impossible to reject any hypothesis or a set that is so small that it does not provide sufficient control of type I errors. Kirk (1995) stated that a family of tests should consist of those tests that are related in terms of their content and intended use. Numerous methods have been proposed for dealing with this problem but no one solution will be acceptable for all situations. One of such methods is the Bonferroni multiplicity correction. It is considered the simplest and most conservative method to control the *familywise error rate*. The correction is based on the idea that if an experimenter is testing n independent or dependent hypotheses on a set of data, then one way of maintaining the *familywise error rate* is to test each individual hypothesis at a statistical significance level of $1/n$ times what it would be if only one hypothesis were tested. So, if we want to keep the *familywise error rate* bounded by α, then the Bonferroni correction is to divide the acceptable α - level by the number of comparisons or hypotheses, which in our case is n.

2.4 Lagrange Multiplier test and Information matrix

The Fisher Information (sometimes referred as information) can be defined as the variance of the score or as the expected value of the observed Information. It describes the amount of information data provide about an unknown parameter. The role of the Fisher Information in the asymptotic theory of maximum-likelihood estimates was emphasized by the statistician R. A Fisher. The Fisher Information is also used in Bayesian statistics (for example, it is used in calculating Jeffrey's prior). The Fisher information matrix is used to calculate the covariance matrices associated with maximum-likelihood estimates. It can be

used in the formulation of test statistics, such as the Wald Test and the Lagrange Multiplier test.

2.4.1 Definitions

The Fisher Information is a way of measuring the amount of information that an observable random variable X carries about an unknown parameter θ upon which the probability of X depends. Let $X = (X_1, X_2, ..., X_n)$ be a random sample and let the probability function of X, which is also the likelihood function of θ be denoted by $f(X|\theta)$ with parameter vector $\theta = (\theta_1, \theta_2, ...\theta_k)$. The partial derivative with respect to θ of the natural logarithm of the likelihood function is called the score (that is, its expected value) is 0. The second moment is called the Fisher information matrix. The Fisher information matrix, $I_n(\theta)$, of sample size n is given by the $k \times k$ symmetric matrix whose i, j_{th} entry is given by the covariance between the first partial derivatives of the log-likelihood,

$$I_n(\theta)_{i,j} = cov\left[\frac{\partial}{\partial\theta_i} lnf(X|\theta), \frac{\partial}{\partial\theta_j} lnf(X|\theta)\right]$$

If $lnf(X|\theta)$ is twice differentiable with respect to θ, then an alternative but equivalent definition for the Fisher information matrix is based on the expected values of the second partial derivatives, and is given by

$$I_n(\theta)_{i,j} = -E\left[\frac{\partial^2}{\partial\theta_i\theta_j} lnf(X|\theta)\right]$$

2.4.2 Lagrange Multiplier test

Rao's score test (often known as the Lagrange Multiplier test) is a statistical test of a simple null hypothesis that a parameter of interest θ is equal to some particular value θ_0. The main advantage of the score test is that it does not require an estimate of the information under the alternative hypothesis or unconstrained maximum likelihood.

Single parameter test

Let L be the likelihood function which depends on univariate parameter θ and let x be the data. The score is $U(\theta)$ where $U(\theta) = \frac{\partial logL(\theta|x)}{\partial \theta}$. The observed information is $I_n(\theta) = -E\left[\frac{\partial^2}{\partial \theta^2} logf(X;\theta)|\theta\right]$. The statistic to test $H_0 : \theta = \theta_0$ is $S(\theta_0) = \frac{U(\theta_0)^2}{I(\theta_0}$ which has an asymptotic distribution of χ_1^2, when H_0 is true.

Relationship with other tests

The likelihood ratio test, the Wald Test and the Score test are asymptotically equivalent tests of hypotheses. All the three use the likelihood of the models being compared to assess their fit. The likelihood is the probability the data given the parameter estimates. The goal of a model is to find values for the parameters (coefficients) that maximize value of the likelihood function, that is, to find the set of parameter estimates that make the data most likely. When testing nested models, the statistics for each test converge to a chi-squared distribution with degrees of freedom equal to the difference in degrees of freedom in the two models.

2.5 Polychoric correlation

In statistics, polychoric correlation is a technique for estimating the correlation between two theorized normally distributed continuous latent variables, from two observed ordinal variables. The polychoric correlation coefficient, introduced by Pearson, is an alternative to the Pearson r specifically for situations in which the variables of interest are continuous but the measurement instruments yield data that may only be ordinal (Pearson and Pearson 1992; Ritchie 1918). In simple terms, it is a measure of the association between two ordinal variables. In other words, it is based on the assumption that the two latent bivariate normally distributed random variables generate couples of ordinal scores. Categories of the two ordinal variables correspond to intervals of corresponding continuous variables. Thus,

measuring the association between ordinal variables means estimating the product moment correlation between the underlying normal variables (Olsson,1979). Babakus (1985) and Olsson (1979b) have shown that the polychoric correlation coefficient, calculated from ordinal transformations of bivariate normal variables, produces an unbiased estimate of the correlation between the original bivariate normal variables.

Furthermore, Babakus offered some evidence supporting the use of polychoric correlation coefficient instead of the Pearson r in maximum likelihood confirmatory factor analysis. Babakus performed a simulation based on a confirmatory factor model with one latent factor and four indicators. Ordinal data were used to calculate four different coefficients; the Pearson r, the polychoric, Spearman's ρ and Kendall's τ. Babakus found that the polychoric's performance was superior. The polychoric correlation coefficient produced parameter estimates that were essentially unbiased and that had the smallest mean squared error. The results were also relatively robust to variations in the shape of the distribution of the ordinal data. Besides its desirable properties, there were some shortcomings. It overestimated both the standard errors of the parameter estimates and the chi-square goodness-of-fit statistic. This overestimation was theorized to be partly due to the use of a maximum likelihood fitting function.

Browne (1982) showed that the maximum likelihood fitting function which is frequently used in causal modeling is actually only one of a class of fitting functions all of which are at least equivalent to the form

$$F = (S - \Sigma)W^{-1}(S - \Sigma)$$

where S and Σ are vectors composed of the $\frac{p(p+1)}{2}$ distinct elements of the empirical and model-based estimates of the covariance matrix of p indicators in the model and W is a matrix of weights of dimension $\frac{p(p+1)}{2} \times \frac{p(p+1)}{2}$. It is the choice of W, the weight matrix, that distinguishes the different functions in this class.

Joreskog and Borbom's Generalized Least Squares (GLS) fitting function is closely

related to the ML and is thought to yield comparable results, especially when there is a good fit between the model and the data. (Note: generalized least squares, as used by Joreskog and Sorbom, should not be confused with generalised least squares methods, as used by Browne 1982 and Muthen 1984. The GLS fitting function discussed here is a special case of the generalized least squares methods of Browne and Muthen). The simplest alternative in this class involves calculating W as an identity matrix. Joreskog and Borbom's GLS is equivalent to that procedure. Browne showed that when W is an estimate of the asymptotic covariance matrix of the covariance matrix of the indicators, the indicators need not be multivariate normal. Several causal modeling software packages include fitting functions based on this finding. Joreskog and Borbom (1985) labeled their version "weighted least squares" (WLS) and they further stated that their version can accommodate ordinal data so long as 1)the sample size is sufficiently large, and 2) the input covariance matrix is composed of polychoric correlation coefficients.

Chapter 3

Applied Techniques and Methodology

3.1 Structural Equation Models

Structural equation modeling (SEM) is a very general, chiefly linear, chiefly cross-sectional statistical modeling technique used to establish relationships between variables. It is an extension of Multiple Linear Regression Analysis (MLRA) and made up of a series of MLRA equations with all equations being fitted simultaneously. SEM approaches encompass diverse statistical techniques including ANOVA/MANOVA, MLRA, Path Analysis, CFA. The basic principle of SEM lies in testing whether or not a specified model, which is theory-based, fits the data.

Assumptions underlying SEM

- Multivariate Normal Distribution of Indicators : Each indicator (observed variable) should be normally distributed for each value of the other indicator. Multivariate normality is required by the Maximum Likelihood Estimation (MLE), which is the predominant method in SEM for estimating structural (path) coefficients. However, Kline (1998) suggests that under conditions of severe non-normality of data, SEM parameter estimates (e.g path estimates) are still fairly accurate but corresponding significant indices are too high.

- Multivariate Normal Distribution of the Latent Dependent Variables : Each dependent latent variable in the model should be normally distributed for each value of each other latent variable. Dichotomous latent variables violate this assumption

- Linearity : SEM assumes linear relationships between latent variables and indicators. However, as it is with regression analysis, it is possible to add logarithmic or other non-linear transformations of the original variable to the model
- Indirect measurement : Typically, all the variables in the model are latent variables
- Multiple Indicators : A minimum of three indicators should be used to measure each latent variable in the model. Multiple indicators are part of a strategy to lower measurement error and increase data reliability.
- Not theoretically underidentified or just identified models : A model is just identified or saturated if there are as many parameters to be estimated as there are elements in the covariance matrix (df =0). Researchers seek an overidentified model, that is , one where the number of observed variances and covariances is greater than the number of parameters to be estimated (df $>$ 0).
- Not empirically identified due to high multicollinearity : A model can be theoretically identified but still not solvable due to empirical problems such as high multicollinearity or path estimates close to zero in non-recursive models.
- Multicollinearity : Complete multicollinearity is assumed to be absent, but correlations among the independents may be modeled explicitly in SEM

Components of a Structural Equation Model

Structural Equation models involves the evaluation of two models: a structural model and a measurement model. The structural model deals with the relationships between latent variables only. The measurement model is the part that relates measured (observed) variables to latent(unobserved) variables. Even though no variables may have been manipulated, variables and factors in SEM may be classified as “independent variables” or “dependent variables”. Such classification is made on the basis of a theoretical causal model, formal

or informal. In SEM, a variable can serve both as a source variable (called an exogenous variable which is analogous to an independent variable) and a result variable (called an endogenous variable which is analogous to a dependent variable). The causal model is presented in a diagram where the names of measured variables(indicators) are within rectangles or squares and the names of factors (latent variables) in ellipses or circles. Rectangles and ellipses are connected with lines having an arrowhead on one (unidirectional causation) or two (no specification of direction of causality) ends. Dependent or endogenous variables are those that have one-way arrows pointing to them and independent variables are those that do not have.

The diagram below shows an example of a Structural Equation Model.

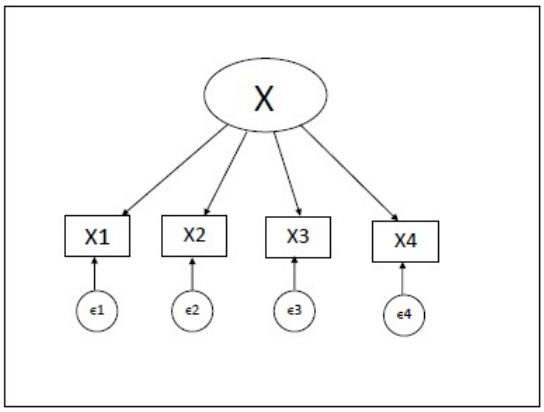

Figure 3.1: SEM example

Then,

$$X1 = \alpha + X\beta_1 + \epsilon_1$$
$$X2 = \alpha + X\beta_2 + \epsilon_2$$
$$X3 = \alpha + X\beta_3 + \epsilon_3$$
$$X4 = \alpha + X\beta_4 + \epsilon_4$$

The variables on the left-hand side of the structural equations are the *endogenous variables* - that is, variables whose values are determined by the model. So, generally, there is one structural equation for each endogenous variable in an SEM. The ϵ's are error variables (that is, $\epsilon_i \sim N(0, \sigma^2)$); they play a role analogous to the error in a single-equation regression model. The remaining variables on the right-hand side of the model are *exogenous variables*, whose values are treated as conditionally fixed; an additional defining characteristic of exogenous variables is that they are assumed to be independent of the errors (much as the predictors in a common regression model are taken to be independent of the error). The α's are structural parameters (regression coefficients) relating the endogenous variables to the exogenous variables and the β's are also structural parameters relating the endogenous variables to one another.

Statistically, the model is evaluated by comparing two variance/covariance matrices. From the data, a sample variance/covariance matrix is calculated. From this matrix and the model, an estimated population variance/covariance matrix is computed. If the estimated population variance/covariance is similar to the known sample variance/covariance, the model is said to fit the data well.

We can, for example, compute the sample variance/covariance matrix of our four (4) observed variables, $Y = (X_1, X_2, X_3, X_4)$. The covariance between any two of our observed

variables, X_i and X_j can be mathematically represented as

$$cov(X_i, X_j) = E\left[(X_i - \mu_i)(X_j - \mu_j)\right]$$

where $\mu_i = E(X_i)$ and $\mu_j = E(X_j)$. Let's look at the variance/covariance matrix of Y.

$$\text{cov (Y)} = \begin{bmatrix} cov(X_1, X_1) & cov(X_1, X2) & cov(X_1, X3) & cov(X_1, X4) \\ cov(X_2, X1) & cov(X_2, X2) & cov(X_2, X3) & cov(X_2, X4) \\ cov(X_3, X1) & cov(X_3, X2) & cov(X_3, X3) & cov(X_3, X4) \\ cov(X_4, X1) & cov(X_4, X2) & cov(X_4, X3) & cov(X_4, X4) \end{bmatrix}$$

where the diagonal of this matrix represents the variance of the vector Y with 4 elements.

Rewriting the matrix, we have $$\text{cov (Y)} = \begin{bmatrix} var(X_1) & cov(X_1, X2) & cov(X_1, X3) & cov(X_1, X4) \\ cov(X_2, X1) & var(X_2) & cov(X_2, X3) & cov(X_2, X4) \\ cov(X_3, X1) & cov(X_3, X2) & var(X_3) & cov(X_3, X4) \\ cov(X_4, X1) & cov(X_4, X2) & cov(X_4, X3) & var(X_4) \end{bmatrix}$$

$$\text{cov (Y)} = \begin{bmatrix} {\beta_1}^2 var(X) + \sigma^2 & \beta_1\beta_2 var(X) + \sigma^2 & \beta_1\beta_3 var(X) + \sigma^2 & \beta_1\beta_4 var(X) + \sigma^2 \\ \beta_2\beta_1 var(X) + \sigma^2 & {\beta_2}^2 varX + \sigma^2 & \beta_2\beta_3 var(X) + \sigma^2 & \beta_2\beta_4 var(X) + \sigma^2 \\ \beta_3\beta_1 var(X) + \sigma^2 & \beta_3\beta_2 var(X) + \sigma^2 & {\beta_3}^2 var(X) + \sigma^2 & \beta_3\beta_4 var(X) + \sigma^2 \\ \beta_4\beta_1 var(X) + \sigma^2 & \beta_4\beta_2 var(X) + \sigma^2 & \beta_4\beta_3 var(X) + \sigma^2 & {\beta_4}^2 var(X) + \sigma^2 \end{bmatrix}$$

This essentially represents the sample variance/covariance matrix.

3.2 Structural Equation Modeling Procedures

The goal of constructing a path diagram or building a structural equation model is to find a model that best fits data (S) to serve as a useful representation of reality and a parsimonious explanation of the data. Latent variable SEM occurs in a series of steps.

The steps include:

- Model Specification
- Model Identification
- Model Estimation
- Testing Model Fit
- Model Respecification (Manipulation)

Model specification is often times considered as the most important step.It involves the formal stating or formulation of the model. This is where primary latent variables and how they relate to each other are specified. It also involves the construction of the measurement model which specifies the relationship between latent and observed variables. This is also the step in which parameters are determined to be $free$ or $fixed$. Fixed parameters are not estimated from the data and are typically fixed to zero (indicating no relationship between variables). Free parameters, however are estimated from the observed data and are believed to be non-zero. Determining which parameters are free or fixed is very crucial because it determines which parameters will be used to compare the hypothesized diagram with the population variance and covariance matrix in testing the fit of the model. The choice of which parameters are free and which are fixed in a model is up to the researcher. The information needed for specification of the model comes from the researcher and their knowledge of theory and prior research in this area. This choice represents the researcher's a *priori* hypothesis about which pathways in a system are important in the generation of the observed system's relational structure (ex. the observed sample variance and covariance matrix)

Model identification concerns whether a unique value for each of the every free parameter can be obtained from the observed data. When we can, the model is identified and when we

cannot, it is underidentified. To determine whether the model is identified or not, compare the number of data points to the number of parameters to be estimated. Since the input data set is the sample variance/covariance matrix, the number of data points is the number of variances and covariances in that matrix, which can be calculated as

$$\frac{p(p+1)}{2}$$

where p is the number of measured variables. If the number of data points equals the number of parameters to be estimated,then the model is just identified or saturated. Such a model will fit the data perfectly. An underidentified model is one in which there are fewer data points than parameters to be estimated. In such a case, the parameters cannot be estimated and the researcher needs to reduce the number of parameters to be estimated. When the number of data points is greater than the number of parameters to be estimated, then the model is overidentified and the analysis can proceed. All of the above depends on the model choice and the specification of $fixed$, *constrained* and $free$ parameters. A parameter is constrained when it is set equal to another parameter. If a parameter is not identified, then there will be two or more values of the parameters that are equally consistent with the data and the researcher will not be able to choose among them empirically, even if the population data were available. So, knowledge of the identification status of a model is crucial for proper estimation and interpretation of a model. In general, the means (μ), variances and covariances (Σ) of the observed variables are identifiable parameters for virtually all observed variables that are part of a model.

Estimation: In this step, start values (which are chosen by the researcher from prior information or by computer programs) of the free parameters are chosen in order to generate an estimated population covariance matrix, $\Sigma(\theta)$ from the model. The means, variances and covariances of the observed variables should equal the corresponding population means, variances and covariances of the observed variables if the model is valid. The goal of es-

timation is to produce a $\Sigma(\theta)$ that is closest to the observed sample covariance matrix, S , with the residual matrix (the difference between $\Sigma(\theta)$ and S) being minimized. Various methods can be used to generate $\Sigma(\theta)$ and the choice of the method largely depends on the data including sample size and distribution. Most processes are iterative. The general form of the minimization function is F= $(S-\sigma(\theta))'W(\text{S - } \sigma(\theta))$ where S is the vector containing the variances and covariances of the observed variables $\sigma(\theta)$ is the vector containing corresponding variances and covariances as predicted by the model and W is the weight matrix.

The weight matrix, W in the function, corresponds to the estimation method used. It is chosen to minimize F and F(N-1) gives the fitting function, in most cases a chi-squared distributed statistic. The performance of the χ^2 is affected by sample size, error distribution, factor distribution and the assumption that factors and errors are independent(Ullman 1996). Some of the commonly used estimation methods are:

Generalized Least Squares

$\boxed{\text{F}_{GLS} = 1/2\text{tr}[([\text{S -}\Sigma(\theta)]\text{W}^{-1})^2]}$ where, tr = trace operator that takes sum of elements on main diagonal of matrix and W^{-1} = optimal weight matrix that must be selected by the researcher (most common choice is S^{-1});

Maximum likelihood (ML)

$\boxed{\text{F}_{ML} = \log|\Sigma| \text{ - } \log|S| + \text{tr}(\text{S}\Sigma^{-1}) - p}$. In this case, $\text{W} = \Sigma^{-1}$ and p is the number of observed variables.

The above estimators are useful for normally distributed data when factors and errors are independent. For non normally distributed data, the Asymptotically Distribution Free Estimator is useful but it is shown only to work well with sample sizes above 2,500(Ullman 1996). Whatever function is chosen, the desired result of the estimation process is to obtain a fitting function that is close to 0. A fitting function score of 0 implies that the model's estimated covariance matrix and the original sample covariance matrix are equal. For ordinal outcomes, most researchers favor using GLS operating on a polychoric correlation matrix for estimation of the SEM.

Model fit: After estimation, the researcher turns to assessing the fit of the model. Model fit can be seen in two ways. One deals with the component fit of the model and the other deals with the overall fit. The component model fit involves testing individual parameters for statistical significance or testing a group of parameters using for example, the Wald tests. Overall fit is the second way to assess model fit. When the ML estimator and other full information estimators are used, a χ^2 test of overall fit is generally available. For example, for the ML estimator, the test statistic, F_{ML} is $F_{ML}(N-1)$, where F_{ML} is evaluated at the final estimates θ of the parameters. When the assumptions of the ML estimator are satisfied, F_{ML} follows an asymptotic χ^2 distribution with degrees of freedom, $\text{df} = (1/2)p(p+1) - t$, where p is the number of observed variables and t is the number of free parameters estimated in the model. If the ratio between χ^2 and the degrees of freedom is less than three, the model is a good fit (Ullman, 1996). The null hypothesis of this χ^2 test is; $H_0 : \Sigma = \Sigma(\theta)$. A statistically significant test statistic casts doubt on the implied moment structure and the model that gave rise to it. A non-significant test statistic is consistent with the model structure. Practically, $H_0 : \Sigma = \Sigma(\theta)$ is too strict for most models because the test does not tolerate even the slightest misspecification and hence in situations with sufficient statistical power (example,when N is large), the null is almost always rejected. Alternative model assessment methods exist. If the overall or companion fit of a model is inadequate, then there is the need to respecify the model.

Model modification: This refers to revisions of an initial model. If the covariance/variance matrix estimated by the model does not adequately reproduce the sample covariance/variance matrix, hypotheses can be revised or adjusted and the model retested. These adjustments/revisions can range from minor (example, introducing a secondary path) to major (example, changing the number of latent variables and their relationships). In adjusting a model, new pathways are added or original pathways are removed. In other words, parameters are changed from fixed to free or from free to fixed. It is important to note that these

adjustments increase the chance of making a Type 1 error.

Empirical procedures exist that are used for model modification. The common ones are the Lagrange Multiplier Index (LM) and the Wald test. Both of these tests report the change in χ^2 value when pathways are revised or adjusted. The LM asks whether addition of free parameters increases model fitness. In other words, it is an estimate of the decrease in the χ^2 test statistic that would result by freeing just a single parameter at a time, which can be misleading when multiple modifications are required. The LM tests uses the same logic forward stepwise regression. The Wald test, on the other hand, asks whether deletion of free parameters increases model fitness. The Wald test follows the logic of backward stepwise regression.

To adjust for increased type one error rates, Ullman(1996) recommends using a low probability value ($p \leq 0.01$) when adding or removing parameters. Because the order in which parameters are freed can affect the choice of remaining parameters, LM should be applied before the Wald test (MacCullum 1986, cited in Ullman 1996)

3.3 Modification Index

In SEM, we often times discover that our hypothesized model does not fit our data. In such a scenario, revisions or changes have to be made to obtain better fitting models. There are a number of theoretically consistent actions (or steps) that can be taken to improve our model fit. One of those theoretical possibilities is resorting to the use of "modification index" which is also known as the Lagrange Multiplier Test Statistic in searching for better fitting models. These indices estimate how much we can reduce the discrepancy between the theoretical model and the data.

Generally, fitting techniques take the form : Minimize (or maximize) some fit function $f(\theta, \theta_c, x)$ with respect to the free parameters θ and a set of constrained parameters, θ_c, for a given set of data, x. How do we improve our model if it does not fit our data?

Let's consider an example. In confirmatory factor analysis(Joreskog, 1969), the parameters in θ and θ_c may be taken to be all the elements in the matrix of the factor loadings, Λ, the covariance matrix of the factors, Φ, and the covariance matrix for the specific factors, Ψ. The model specifies that the values of some of these parameters are known a priori (e.g., some $\lambda_{ij} = 0$, $\phi_{kk} = 1$ or, all $\psi_{ij} = 0$, $i \neq j$) or equal to some other elements in these matrices (equality constraints). For maximum likelihood, the fit function is

$$f = 2n(log|\Sigma| - log|S| + tr(\Sigma^{-1}S) - p), \tag{3.-1}$$

where S is the sample covariance matrix (a function of the data x), n is the sample size minus one, p is the number of variables and

$$\Sigma = \Lambda\Phi\Lambda^{'} + \Psi, \tag{3.0}$$

the model-implied covariance matrix. If the model does not fit our data, we would want to know which fixed parameter or which equality constraint should be relaxed to obtain maximum possible decrease in the value of the function f.

In the general formulation, relaxing a fixed parameter or an equality constraint simply means moving one parameter from θ_c to θ. The question is : which one should we choose to maximize the reduction in f? For each parameter in θ_c, we shall construct a modification index (MI) which is approximately equal to the decrease in f that will be obtained if this parameter is relaxed.

3.3.1 Construction of Modification Index

Let $g = \frac{\partial f}{\partial \theta}$ be the gradient vector of a fitting function f and let $E = E\left[\frac{\partial^2 f}{\partial \theta \theta'}\right]$ be the matrix of expected second order derivatives of f. The estimates of θ, denoted by $\hat{\theta}$ are those values that minimize $f(\theta)$. Let g and E evaluated at $\theta = \hat{\theta}$ be denoted by $\hat{g}$ and $\hat{E}$ respectively. Hence we know that $\hat{g} = 0$ and that $\hat{E}$ can be assumed to be positive definite

if the estimated model is identified.

Most fit functions are approximately quadratic around the solution $\hat{\theta}$ and we know that, for a sufficiently large sample, the f matrix is an approximation of the second order derivatives of f. Thus, we can approximate f around $\hat{\theta}$ by the Taylor expansion

$$f \approx \hat{f} + (\theta - \hat{\theta})' \hat{g} + \frac{1}{2}(\theta - \hat{\theta})' \hat{f} (\theta - \hat{\theta}) \tag{3.1}$$

where $\hat{f}$ denotes the value of f at the solution, $\hat{\theta}$. Suppose now that we want to add a previously fixed parameter to the vector of free parameters. Let this new free parameter be denoted by θ_1, and the value to which it was fixed be $\hat{\theta_1}$. Then the Taylor expansion is

$$f \approx \hat{f} + \begin{bmatrix} \theta - \hat{\theta} \\ \theta_1 - \hat{\theta_1} \end{bmatrix}' \begin{bmatrix} \hat{g} \\ \hat{g_1} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_1 - \hat{\theta_1} \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{d} \\ \hat{d}' & \hat{k} \end{bmatrix} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_1 - \hat{\theta_1} \end{bmatrix} \tag{3.2}$$

where $\hat{g_1} = \frac{\partial f}{\partial \theta_1}$, $\hat{d}$ is the vector of expected second order derivatives $E\left[\frac{\partial^2 f}{(\partial \theta \theta_1)}\right]$, and $\hat{k}$ is $E\left[\frac{\partial^2 f}{(\partial \theta_1 \theta_1)}\right]$, all first and second derivatives evaluated at $(\hat{\theta}, \hat{\theta_1})$. In order to study how much the function in (3.2) will be decreased, we find new estimates by minimizing f in (3.2) with respect to $(\hat{\theta}, \hat{\theta_1})$. The minimum, $\hat{\hat{f}}$, must satisfy

$$\begin{bmatrix} \frac{\partial f}{\partial \theta} \\ \frac{\partial f}{\partial \theta_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.3}$$

which, since $\hat{g} = 0$, becomes

$$\begin{bmatrix} 0 \\ \hat{g_1} \end{bmatrix} + \begin{bmatrix} \hat{E} & \hat{d} \\ \hat{d}' & \hat{k} \end{bmatrix} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_1 - \hat{\theta_1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.4}$$

substituting (3.4) into (3.2) gives

$$\hat{\hat{f}} \approx \hat{f} - \begin{bmatrix} 0 \\ g_1 \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{d} \\ \hat{d}' & \hat{k} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ g_1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 \\ \hat{g_1} \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{d} \\ \hat{d}' & \hat{k} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \hat{g_1} \end{bmatrix} = \hat{f} - \frac{\frac{1}{2}\hat{g_1}^2}{\hat{k} - \hat{d}'\hat{E}\hat{d}} \tag{3.5}$$

The last step follows by using the formula for the inverse of a partitioned symmetric matrix,

$$\begin{bmatrix} \hat{E} & \hat{d} \\ \hat{d}' & \hat{k} \end{bmatrix}^{-1} = \begin{bmatrix} \hat{E}^{-1} + h(\hat{E}^{-1}\hat{d})(\hat{E}^{-1}\hat{d})' & -h\hat{E}^{-1}\hat{d} \\ -h(\hat{E}^{-1}\hat{d})' & h \end{bmatrix}$$

where $h = \frac{1}{\hat{k} - \hat{d}'\hat{E}^{-1}\hat{d}}$. It has been assumed that also the matrix $\begin{bmatrix} \hat{E} & \hat{d} \\ \hat{d}' & \hat{k} \end{bmatrix}$ is positive definite. This is exprected to be the case if the model with θ_1 added as a free parameter is identified. The decrease in function value is $\hat{f} - \hat{\hat{f}}$. Thus, by defining the modification index as

$$MI = \frac{\frac{1}{2}\hat{g_1}^2}{\hat{k} - \hat{d}'\hat{E}\hat{d}}, \tag{3.6}$$

we have an approximate estimate of how much the fit function will decrease if one adds a parameter θ_1 previously fixed at $\hat{\theta_1}$ to the set of free parameters. The term $\hat{d}'\hat{E}\hat{d}$ in (3.6) can be seen as an adjustment to the MI that is caused by the change in the parameters in θ when θ_1 is added. Since E is positive definite, this term is always greater than zero, and this means that we get an increase in the MI when change in θ is taken into account.

We have so far looked at the case of freeing a fixed parameter. Let's look at what happens when two or more parameters that are constrained to be equal are relaxed, that is, the effect of relaxing such constraints. Let $\theta_{\mathbf{c}}$ now denote a vector of p parameters all of which are constrained to be equal to the parameter θ_2. This is to say that $\theta_{\mathbf{c}} = \theta_2 i$, where i is p-component vector of all ones. In addition, we have, as before a set of free parameters, θ. We want to compute an estimate of how much the fit function will decrease if we let θ_2 be free while retaining the equality constraints for the elements in $\theta_{\mathbf{c}}$. Suppose f has been

minimized with respect to θ and θ_2 yielding $\hat{\theta}$ and $\hat{\theta_2}$ so that $\hat{\theta_{\mathbf{c}}} = \hat{\theta_2 i}$. Using the analogy of (3.2), we have

$$f \approx \hat{f} + \begin{bmatrix} \theta - \hat{\theta} \\ \theta_c - \hat{\theta_c} \\ \theta_2 - \hat{\theta_2} \end{bmatrix}' \begin{bmatrix} \hat{g} \\ \hat{g_c} \\ \hat{g_2} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_c - \hat{\theta_c} \\ \theta_2 - \hat{\theta_2} \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{D_1} & \hat{d_2} \\ \hat{D_1}' & \hat{C_{11}} & \hat{C_{12}} \\ \hat{d_2}' & \hat{C_{21}} & \hat{C_{22}} \end{bmatrix} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_c - \hat{\theta_c} \\ \theta_2 - \hat{\theta_2} \end{bmatrix} \tag{3.7}$$

where $\hat{g}$, $\hat{g_c}$ and $\hat{g_2}$ are first order derivatives of f with respect to θ, θ_c and θ_2. D_1 is the matrix $E\left[\frac{\partial^2 f}{\partial \theta {\theta_c}'}\right]$, $\hat{d_2}$ is the vector $E\left[\frac{\partial^2 f}{\partial \theta \theta_2}\right]$, $\hat{E}$ is the expected second order derivatives for the elements in θ and C is the matrix

$$C = \begin{bmatrix} C_{11} & C12 \\ C_{21} & C_{22} \end{bmatrix} = \begin{bmatrix} E\left[\frac{\partial^2 f}{\partial \theta_c {\theta_c}'}\right] & E\left[\frac{\partial^2 f}{\partial \theta_c \theta_2}\right] \\ E\left[\frac{\partial^2 f}{\partial \theta_2 {\theta_c}'}\right] & E\left[\frac{\partial^2 f}{\partial \theta_2 \theta_2}\right] \end{bmatrix} \tag{3.8}$$

Substituting $\hat{\theta_c}$ with $\hat{\theta_2 i}$ in (3.7) and denoting element in $\theta_{\mathbf{c}}$ by θ_c we get

$$f \approx \hat{f} + \begin{bmatrix} \theta - \hat{\theta} \\ \mathrm{i}(\theta_c - \hat{\theta_2}) \\ \theta_2 - \hat{\theta_2} \end{bmatrix}' \begin{bmatrix} \hat{g} \\ \hat{g_c} \\ \hat{g_2} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \theta - \hat{\theta} \\ \mathrm{i}(\theta_c - \hat{\theta_2}) \\ \theta_2 - \hat{\theta_2} \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{D_1} & \hat{d_2} \\ \hat{D_1}' & \hat{C_{11}} & \hat{C_{12}} \\ \hat{d_2}' & \hat{C_{21}} & \hat{C_{22}} \end{bmatrix} \begin{bmatrix} \theta - \hat{\theta} \\ \mathrm{i}(\theta_c - \hat{\theta_2}) \\ \theta_2 - \hat{\theta_2} \end{bmatrix}$$

By the constraint $\hat{\theta_c} = \hat{\theta_2 i}$ we know that $i' \hat{g_c} + \hat{g_2} = 0$. Thus, by replacing $i' \hat{g_c}$ with $-\hat{g_2}$ in (3.7) and utilizing the fact that $g = 0$ we get

$$f \approx \hat{f} + \begin{bmatrix} \theta - \hat{\theta} \\ \theta_c - \hat{\theta_2} \\ \theta_2 - \hat{\theta_2} \end{bmatrix}' \begin{bmatrix} \hat{0} \\ -\hat{g_2} \\ \hat{g_2} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_c - \hat{\theta_2} \\ \theta_2 - \hat{\theta_2} \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{D_1 i} & \hat{d_2} \\ i' \hat{D_1}' & i' \hat{C_{11}} i & i' \hat{C_{12}} \\ \hat{d_2}' & \hat{C_{21}} i & \hat{C_{22}} \end{bmatrix} \begin{bmatrix} \theta - \hat{\theta} \\ \theta_c - \hat{\theta_2} \\ \theta_2 - \hat{\theta_2} \end{bmatrix}$$

Minimizing f with respect to θ, θ_c, and θ_2, results in an expression for the MI as

$$MI = \frac{1}{2}\hat{g_2}^2 \begin{bmatrix} \hat{0} \\ -1 \\ 1 \end{bmatrix}' \begin{bmatrix} \hat{E} & \hat{D_1}i & \hat{d_2} \\ i'\hat{D_1}' & i'\hat{C_{11}}i & i'\hat{C_{12}} \\ \hat{d_2}' & \hat{C_{21}}i & \hat{C_{22}} \end{bmatrix}^{-1} \begin{bmatrix} \hat{0} \\ -1 \\ 1 \end{bmatrix} \tag{3.9}$$

By the use of the formula for inversion of a partitioned matrix, (3.9) can be written

$$MI = \frac{1}{2}\hat{g_2}^2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}' \left[\begin{bmatrix} i'\hat{C_{11}}i & i'\hat{C_{12}} \\ \hat{C_{21}}i & \hat{C_{22}} \end{bmatrix} - \begin{bmatrix} i'\hat{D_1'} \\ \hat{d_2'} \end{bmatrix} \hat{E^{-1}}(\hat{D_1}i\hat{d_2}) \right]^{-1} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \tag{3.10}$$

Thus, denoting the elements of the matrix to be inverted in (3.10) by m_{ij} we can write the MI as

$$MI = \frac{\frac{1}{2}\hat{g_2}^2(m_{11} + 2m_{21} + m_{22})}{m_{11}m_{22} - {m_{12}}^2} \tag{3.11}$$

The inverse of $\hat{E}$ in (3.10) can be computed from the estimate of the inverse of the expected second order derivatives at the constrained solution. Let the second order derivatives at the solution be denoted

$$\hat{E^*} = \begin{bmatrix} E & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

where E, as before, is the matrix of expected second order derivatives for the free parameters and

$$e_{12} = e'_{21} = \hat{D_1}i + \hat{d_2}$$

$$e_{22} = i'C_{11}i + i'C_{12} + C_{21}i + c_{22}.$$

Thus, by using the formula for the inverse of a partitioned matrix we can get

$$\hat{E}^{-1} = E^{11} - \frac{e^{12}e^{21}}{e^{22}},$$

where the superscripts denote the matrices in the inverse of the partitioned matrix E^*.

If the model is correct and the observed variables have a multivariate normal distribution, then our maximum likelihood fit function is approximately distributed in large samples as central χ^2 with degrees of freedom equal to $\frac{p(p+1)}{2}$ - t, where t is the number of free parameters of the model and p is the number of observed variables. The MIs will then be approximately distributed as χ^2 with 1 degree of freedom, which also holds when the GLS fit function is used.

The procedure described above is implemented in computer programs (Joreskog and Sorbom, 1984). This program computes the estimates of the parameters of a model by an iterative minimization of the fit function f. In order to compute standard errors of the parameter estimates, the information matrix, that is, the matrix E, and its inverse are evaluated at the minimum. Thus, to compute the MIs, we need only compute the vector d and the scalar k in (3.4) for each fixed parameter of the model. In addition, we will get estimates of the parameters of the modified model by the solution of (3.2). This has been subsumed in the computer program by a procedure for automatic modification of a model. The program finds the parameter with the largest MI, adds this to the set of free parameters and goes on repeatedly until none of the MIs is greater than a specified value. Since the MIs are approximately χ^2 with 1 degree of freedom, the value could be chosen such that the procedure stops when there is no MI significant at, for example, the 1%-level. Or, in other words, the procedure stops when no significantly better model can be found. At each step, the estimates from the solution of (3.2) in the previous step are used as starting values for the iterative minimization of the fit function, f. This procedure has been found to work very efficiently, at least numerically. How it works from a substantial point of view must however be judged from case to case.

For a fixed parameter which would be nonidentified if set free, the first order derivative must be zero. If this was not the case, it would mean that we could estimate the model with the parameter free and obtain a decrease in the fit function, since the matrix of expected second order derivatives is positive definite. This contradicts the non-identification status of the parameter, since changing a non-identified parameter can not change the value of the fit function. Thus, by (3.4), the MI is zero for each of such parameter. This in turn means that, as long as we start with an identified model, the procedure for automatic modification should produce identified models.

3.4 Type I Error Control Using Scheffe

Although researchers pay a great deal of attention to the control of Type I error in ANOVA-type models, the issue of Type I error control has received considerably less attention in the SEM literature. Users of SEM often ignore this principle, but controlling type I errors across multiple tests of individual parameters is vital in SEM. Failure to control type I error rate can result in freeing parameters that may not actually be statistically significantly different from their fixed values in the population. This practice results in less parsimonious models, affecting the researchers' ability to draw valid conclusions (Green and Babyak, 1997). From an ANOVA perspective, an F test-statistic (F_i) could be constructed for each of the infinite possible post hoc contrasts over k sample means. Some of the contrasts might produce highly correlated test-statistics while others may not. Type I error control procedures suggested by Bonferroni, Dunn- Sidak (Dunn, 1958; Sidak, 1967), or Studentized maximum modulus (Roy and Bose, 1953; Tukey, 1953) do not take into account the correlational structure of the contrasts and so divide alpha among the infinite number of possible post hoc contrasts. From the outset, we might not even know the number of possible post hoc contrasts so if it's very large, then dividing alpha by this large number might be approximately zero. This usually yields an infinitely large critical F value

for use in post hoc testing.

However, Scheffe's method of maintaining a familywise type I error control across all possible contrasts takes into account the correlational structure of the contrasts. Scheffe used a multivariate F distribution (of infinite dimension) to select a critical value for all post hoc testing. Particularly, for the infinite family of post hoc contrast null hypotheses $H_1, ..., H_\infty$, the omnibus null which is the null of the intersection

$$H_0 = \bigcap H_i | i = 1, ..., \infty,$$

alpha-level control over familywise type I error rate is obtained by choosing a critical value F_* such that

$$Pr[\cap(F_i \leq F_*); i = 1, ..., \infty | H_0] = 1 - \alpha.$$

This choice of critical value is equivalent to that F_* for which

$$Pr[max(F_i) > F_* | i = 1, ..., \infty, | H_0] = \alpha$$

Thus, controlling the probability that one or more test statistics of the contrasts exceed their critical value is identical to controlling the probability that the largest of these exceeds the critical value. The largest of these contrasts is known and it is the contrast that captures all the between-group variability. It is the contrast whose sum of squares is exactly $SS_{Between}$; no other contrast account for more variability than this one. So this contrast (largest contrast) will have one *df* and its observed F is given by $(SS_B/1)/MS_{Within}$. which is always $(k-1)$ times larger than the omnibus F test statistic. For this reason, Scheffe proposed a critical value such that it is $(k-1)$ times larger than that required in the omnibus test,

$$F_* = (k-1)(F_{k-1,N-k})$$

This controls the probability that one or more post hoc contrasts will exceed F_* to exactly the desired alpha level under the complete null hypothesis.

Applying these concepts in ANOVA is similar to applying them in the SEM setting. In the SEM setting, we assume that there are c possible post hoc tests of parameters that could be conducted after imposing an initial hypothesized model. The value c can be extremely large and increases as the complexity of the model increases. If these finite c possible univariate post hoc model modifications could be enumerated and an observed χ_i^2 improvement statistic with 1 *df* be determined for each, then some of the modifications would be partially redundant (that is, account for common portions of the total badness of fit in the model χ^2) while others could yield completely independent decreases in the overall badness of fit. Given that, the initially hypothesized model is indeed correct in the population (that is, no post hoc respecifications represent true population relation) and that the distributional assumptions is satisfied, alpha-level control over familywise respecification type I error rate may be obtained by choosing χ_*^2 such that

$$Pr[\cap(\chi_i^2 \leq \chi_*^2)|i = 1, ..., c] = 1 - \alpha$$

This is equivalent to that χ_*^2 for which

$$Pr[max(\chi_i^2) > \chi_*^2; i = 1, ..., c] = \alpha$$

In ANOVA, we know the largest contrast's test statistic and can choose a critical value to exact precise post hoc type I error. However, this is not the case in SEM because the value of χ_*^2 to exact precise error rate control varies as a function of the number of, and interrelations among, all possible post hoc modifications. Deriving such a critical value is though not insurmountable but will require a numerical approximation procedure. A slightly conservative but relatively simpler strategy is to treat the SEM scenario like that of the ANOVA. That is, if there is a single model respecification that is able to explain all the lack of fit in the model chi-square, then we use an alpha-level critical chi-square from

the central distribution with df_{model} to exact precise alpha-level type I error control over all potential univariate respecifications.

Using this critical value for all univariate model modification tests constitute a simultaneous testing procedure and as it has been stated that, it is a conservative strategy. To alleviate some of the conservatism, we employ a sequential testing procedure. Kaplan and Wenger (1993) recommended that model modification is most effective in univariate increments. Since parameters are often asymptotically dependent to some extent, the decision to incorporate a new parameter has consequences for other modifications still under consideration. In this framework, adding a new parameter to the model yields a new model with one fewer degree of freedom than the initial model. Thus, a slightly smaller set of potential modifications remains and a new critical value based on a chi-square distribution with one fewer degree of freedom may be used at this stage. of the modification procedure. This sequential procedure continues until no more modifications are of theoretical interest or until the critical value at that stage of the testing cannot be exceeded.

Chapter 4

Data Analysis

4.1 Obtaining the data

The data set used in the analysis was obtained from the UDSMR software portal using SQL queries. Due to some of the policies of the hospital, information such as the treating physician and other personal information were not made available to us. However, the retrospective data set contained 107 variables and 1715 observations. We were only interested in patients classified as orthopedic or stroke. The orthopedic data was made up of 805 observations whiles the stroke data contained 286 observations.

4.2 Data coding in R

Since our primary interest is to investigate how the FIM model fits our data using SEM, we used procedures for building a model in SEM. The diagram of the structure of the FIM was used in building our model.

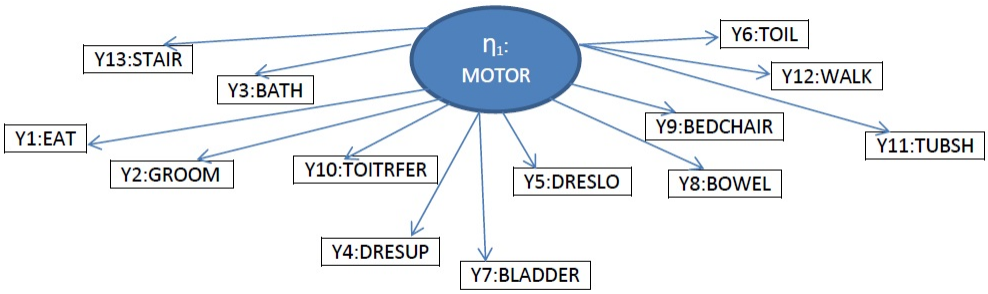

Figure 4.1: Path diagram for the MOTOR dimension of FIM

.

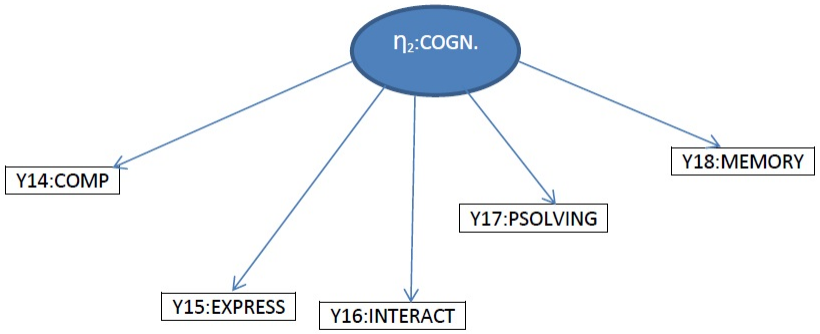

Figure 4.2: Path Diagram for the COGNITIVE dimension of FIM

.

Directed arrows designate regression coefficients and bidirectional arrows signify covariances. The covariances represented by the two bidirectional arrows with broken lines are not included in an initial model specified for our data. The directional arrows are labeled with Greek Letters, that is, $\theta_1 - \theta_{36}$ (orthopedic) and $\lambda_1 - \lambda_{36}$ (stroke) and they represent the corresponding regression coefficients.

Cognitive and Motor functions are our two latent variables with data collected on stroke and orthopedic patients. Our model does not have a structural submodel because no relationship (no correlation) is allowed between the latent variables. However, we have a measurement submodel with equations for the eighteen (18) indicators of our two (2) latent

variables.

Internally, the *sem* function, which is used to fit structural equation models in R, employs the recticular action model (RAM) formulation of the model due to McArdle (1980) and McArdle and McDonald (1984).

In the RAM model, the vectors u and v are related by

$$v = Av + u$$

where the vector v contains indicator variables, directly observed exogenous variables, and latent exogenous and endogenous variables while u (which may overlap with v) contains directly observed and latent exogenous variables, measurement-error variables and structural-error variables. All of these variables might not be present in every model, for example, our FIM model does not contain directly observed exogenous variables, measured-error variables and structural-error variables. The matrix A contains regression coefficients (both structural parameters and factor loadings). As is typically the case, most of the entries of A are prespecified to be 0 whereas others are set to 1.

In the RAM formulation, the matrix P which is

$$P = E(uu')$$

contains covariances among the elements of u (assuming that the elements of u have zero means). One advantage of the RAM formulation of the structural equation model is that the elements of the A and P matrices can be read off the path diagram for the model, with single-headed arrows corresponding to elements of A and double-headed arrows to elements of P, taking into account the fact that variances (as opposed to covariances) of exogenous variables and errors do not appear directly in the path diagram. To make the variances explicit, they are shown as self-directed double-headed arrows.

Model specification in the *sem* package is handled most conveniently via the *specifymodel* function :

The line-number prompts (refer to Appendix A) are supplied by "specifyModel". There are three entries in each line, separated by commas. A single-headed arrow in the first entry indicates a regression coefficient and corresponds to a single-headed arrow in the path diagram; likewise a double-headed arrow represents a variance or covariance and corresponds to a double-headed arrow. The second entry in each line gives the (arbitrary) name of a free parameter to be estimated

Entering the name "NA"(missing) indicates that a parameter is to be fixed to a particular value. The third entry in each line either assigns a value to a fixed parameter or sets a start value for a free parameter ; in the latter case, entering "NA" causes *sem* to pick the start value.

To estimate the model, the covariance or raw-moment matrix among the observed variables has to be computed. In our FIM data, we use the *hetcor*(*polycor*) function in R which computes a heterogenous correlation matrix , consisting of Pearson-product moment correlations between numeric variables, polyserial correlations between numeric and ordinal variables, and polychoric correlations between ordinal variables. We achieved this by (Refer to first four lines of Appendix D) and to estimate the model (refer to the fifth line of Appendix D)

The first argument to *sem* is the model-specification object returned by the specifymodel. The second argument r.FIM, is the input covariance matrix. The third argument is the number of observations on which the covariances are based.The SEM package in R was used in building our model. However, the MLE estimation of the SEM requires multivariate normality. Since the FIM scores violate multivariate normality, the estimation method GLS was used.

As is typical of R programs, *sem* returns an object rather than a printed report. We performed additional computations on *sem* objects, for example, producing modification indices.

```
> modIndices(sem.FIMST)

 5 largest modification indices, A matrix:
       COGSTROKE<-EATSTR   COGSTROKE<-MOTORSTROKE   MOTORSTROKE<-COGSTROKE
                73.45967                  68.77739                  68.77739
MOTORSTROKE<-INTERACTSTR  MOTORSTROKE<-EXPRESSSTR
                67.32746                  63.57239

  5 largest modification indices, P matrix:
    MOTORSTROKE<->COGSTROKE     DRESSLOSTR<->DRESSUPSTR
                   68.77739                    42.43594
     TUBSHOWERSTR<->BATHSTR TUBSHOWERSTR<->TOITRSFERSTR
                   42.07205                    33.83882
         COGSTROKE<->EATSTR
                   22.59519
```

The modification indices are a one-df-chi-square score ("Lagrange multiplier") test statistics for the fixed and constrained parameters in a structural equation model. They may be regarded as an estimate of the improvement in the likelihood-ratio chi-square statistic for the model if the corresponding parameter is respecified as a free parameter. The modIndices function also estimates the change in the value of a fixed or constrained parameter if the parameter is respecified as free.

4.3 Results

Below is the output of the original FIM model for the ORTHOPEDIC PATIENTS

```
Model Chisquare =  933.97   Df =  135 Pr(>Chisq) = 1.3098e-119
 Goodness-of-fit index =  0.8845
 Adjusted goodness-of-fit index =  0.8537
 RMSEA index =  0.085797
 Bentler CFI =  0.88662
 SRMR =  0.17916
 AIC =  1006
 BIC =  1174.8
```

The output (refer to Appendix B) depicts that the solution converges in 51 iterations with a model χ^2 value of 933.97 and 135 degrees of freedom. The degrees of freedom generally is the number of non-redundant elements of the $p \times p$ covariance minus the number of free parameters. In our case, the covariance matrix is 18×18, so there will be

$$\frac{p(p+1)}{2} = \frac{18(19)}{2} = 171$$

non -redundant elements. Since there 36 free parameters $(\theta_1 - \theta_{36})$, there are 171 - 36 = 135 degrees of freedom.

The χ^2 value has a $p - value$ that is very small, and so the null hypothesis of perfect fit is rejected. The output also includes parameter values, estimates of their standard errors and asymptotically normal statistics testing the hypothesis that the parameter value is zero in the population.

4.4 Evaluating and Improving Model fit

One input that is necessary when assessing a model is the examination of some fit statistics. The fact that the hypothesis of perfect fit is rejected is,in itself not informative. The more relevant questions may be 1)how bad is the misfit, and 2)how precisely or accurate do we determine the degree or level of misfit. In regression modeling, for example, we normally report the R^2 which is a descriptive index used in evaluating goodness-of-fit. Use of the R^2 is however subjective because there are no fixed guidelines or rule of thumb for it and so $R^2 = 0.50$ may be good sometimes, $R^2 = 0.30$ may also be good etc. In SEM, the issue of model evaluation explodes because there are a plethora of fit indices. These indices vary in relation to sample size, parsimony, absolute fit or relative fit. In all, these different indices complement each other. Gerbing and Anderson (1992) describe the situation as being analogous to the difficulty in answering the question, "what's the best car on the market"?. The answer is that there is no best car. The definition of "best" depends on the objective: do you wish to drive a fast car, stylish car or a safe care? It has been agreed that SEM modelers or researches can report the following profile of indices: the χ^2 (and its degrees of freedom and $p-value$), the standardized root mean square residual (SRMR), the root mean square error of approximation (RMSEA) and the comparative fit index (CFI). Ideally, for a model that fits the data, the χ^2 would not be significant ($p \leq 0.05$), the SRMR would be close to 0.09 (or lower, Hu and Bentler, 1999). All the SEM fit indices are descriptive statistics with the exception of the χ^2 which is inferential. This means statements regarding significance of hypothesis testing may be made only for the χ^2 and for the others,there are only guidelines or "rules of thumb" for assessing goodness-of-fit. This trait make it seem like it is the only statistic that must be reported. However, the χ^2, like any other statistics has its own problems. It is sensitive to sample size (Gerbing and Anderson, 1985). It is important to have a large sample so that parameters can be estimated precisely but as N increases, χ^2 blows up. A χ^2 will always be significant (indicating a poor fit) even with modest sample size. It has therefore been suggested that

a model demonstrates reasonable fit if the statistic adjusted by its degrees of freedom does not exceed 3.0 (Kline, 2004 cited in Iacobucci, 2009): $\chi^2/\text{ df} \leq 3$

4.5 Updating the model

Joreskog introduced the approach which states that a confirmatory model based on theory can be started with and then updated. The update is done by the addition of factor loadings with the help of "modification indices". These indices attempt to estimate which missing paths, if added to the current model, would result in the greatest reduction of the χ^2 fit statistic. The modification indices can be obtained from SEM as follows

```
5 largest modification indices, A matrix:
  COGORTHO<-MOTORORTHO   MOTORORTHO<-COGORTHO MOTORORTHO<-PSOLVINGOT
             203.4893               203.4893               192.9853
 MOTORORTHO<-EXPRESSOT MOTORORTHO<-INTERACTOT
             175.6810               172.8639

  5 largest modification indices, P matrix:
MOTORORTHO<->COGORTHO    EXPRESSOT<->COMPOT  TUBSHOWEROT<->BATHOT
            203.48930             104.40718              77.18015
     COGORTHO<->EATOT   BOWELOT<->BLADDEROT
             51.72312              43.90106
```

An entry in the A matrix is of the form,

```
<endogenous variable> : <exogenous variable>
```

So, the largest modification index (MI), labeled

```
COGORTHO:MOTORORTHO
```

indicates that, the model χ^2 fit index would decrease by approximately 203.49 if a path from MOTORORTHO to COGORTHO is added to the original model. A path from MOTORORTHO to COGORTHO, however, cannot be added because it conflicts or it is not consistent with the theory behind the model, which states that every indicator must load onto a single latent variable (cross loadings are not permitted). Besides, by allowing that path, we are creating a new regression between MOTORORTHO and COGORTHO abilities which is theoretically inconsistent.

We now resort to using modification indices in the P matrix. The largest MI in the P is labeled MOTORORTHO:COGORTHO, that is the latent variables MOTOR and COGNITION should be allowed to covary. So, we add that path to our model and we have the following results.

```
Model Chisquare =  696.27   Df =  134 Pr(>Chisq) = 8.2608e-77
Goodness-of-fit index =  0.90422
Adjusted goodness-of-fit index =  0.87777
RMSEA index =  0.072242
Bentler CFI =  0.92021
SRMR =  0.057517
AIC =  770.27
BIC =  943.83
```

```
the37 0.56484  0.028012  20.1637  2.0405e-90 COGORTHO <--> MOTORORTHO
```

We observe that the new parameter, θ_{37} has an estimate value of 0.56 and is highly significant. Besides, the model χ^2 statistic decreased to 696.27, a 237 point decrease even greater than what was predicted by our MI. According to statistical theory in Steiger, Shapiro and Browne (1984), since the models are nested (the original model is a special case of the revised model, where one of the parameters is constrained to zero), the difference in the two χ^2 values can be treated as a χ^2 with degrees of freedom equal to the difference in their degrees of freedom (that is, 1 in this case). The resulting chi-square difference statistic provides a method for testing whether there is a statistically different improvement in fit. Obviously, there is. However, the RMSEA statistic for the improved model is still large (0.07) than the recommended value. We recompute MIs, that is, we compute a new set of MIs based on the new (or improved) model. So, whenever we make a modification or change to our model, new MIs are computed based on the most current model.

```
> modIndices(sem.FIM)


 5 largest modification indices, A matrix:
  COMPOT<-EXPRESSOT   EXPRESSOT<-COMPOT BATHOT<-TUBSHOWEROT TUBSHOWEROT<-BATHOT
          109.74834           109.74643            83.13079             83.13023
    EATOT<-COGORTHO
           49.13037


  5 largest modification indices, P matrix:
  EXPRESSOT<->COMPOT TUBSHOWEROT<->BATHOT     COGORTHO<->EATOT
           109.74742             83.13037             49.13023
  MOTORORTHO<->EATOT  BOWELOT<->BLADDEROT
            49.12994             39.02083
```

The largest MI is for a covariation between EXPRESSOT and COMPOT. Adding this path to the model gives:

```
Model Chisquare =  600.52   Df =  133 Pr(>Chisq)=1.4292e-60
Goodness-of-fit index =  0.91678
Adjusted goodness-of-fit index =  0.893
RMSEA index =  0.066122
Bentler CFI =  0.93366
SRMR =  0.05709
AIC =  676.52
BIC =  854.78
```

```
the38 0.09083  0.011175   8.1277  4.3761e-16 COMPOT <--> EXPRESSOT
```

The new parameter, θ_{38}, has a value of 0.09 which is significant. The model χ^2 dropped to 600.52 (the chi square difference is significant) and the RMSEA, at 0.066. We can continue to examine MIs, but we may be reaching a point of diminishing returns where we are capitalizing on chance.

```
> modIndices(sem.FIM)

 5 largest modification indices, A matrix:
TUBSHOWEROT<-BATHOT BATHOT<-TUBSHOWEROT    EATOT<-EXPRESSOT     EATOT<-COGORTHO
          83.16411            83.16402            47.61053            46.58924
    COGORTHO<-EATOT
          46.58797
```

```
 5 largest modification indices, P matrix:
TUBSHOWEROT<->BATHOT      COGORTHO<->EATOT    MOTORORTHO<->EATOT
            83.16469              46.58853              46.58709
 BOWELOT<->BLADDEROT TOITRSFEROT<->GROOMOT
            39.06080              33.19736
```

Now, we can see how all the modification indices are rather close, and are now much smaller than before. Let's allow TUBSHOWEROT and BATHOT to covary and see what happens:

```
Model Chisquare =  518   Df =  132 Pr(>Chisq) = 3.8982e-47
 Goodness-of-fit index =  0.92686
 Adjusted goodness-of-fit index =  0.90525
 RMSEA index =  0.060308
 Bentler CFI =  0.94523
 SRMR =  0.053662
 AIC =  596
 BIC =  778.94
```

```
the39 0.231252 0.028437   8.1322  4.2157e-16 BATHOT <--> TUBSHOWEROT
```

Many of these modifications will be statistically significant at $\alpha = 0.05$ when considered alone (i.e 1 df). Under our proposed method, a modification would need to exceed a critical chi-square value with 135 df, or 163.12 (assuming $\alpha = 0.05$) , for that modification to be considered.

Let's examine the ten largest modification indices (MIs) below

Table 4.1: MIs for ORTHOPEDIC PATIENTS

Modification	χ_2-value	P-value
COG $\leftarrow$ MOTOR	203.4893	0.000
MOTOR $\leftarrow$ COG	203.4893	0.000
COG $\leftrightarrow$ MOTOR	203.4893	0.000
MOTOR $\leftarrow$ PSOLVING	192.9853	0.000
MOTOR $\leftarrow$ EXPRESS	175.6810	0.000
MOTOR $\leftarrow$ INTERACT	172.6839	0.000
EXPRESS $\leftrightarrow$ COMP	104.4072	0.000
TUBSH $\leftrightarrow$ BATH	77.18015	0.000
COG $\leftrightarrow$ EAT	51.72312	0.000
BOWEL $\leftrightarrow$ BLADDER	43.90106	0.000

As can be seen, all the MIs are statistically significant at $\alpha = 0.05$. Notice that the first six indices exceed our proposed critical chi-square value of 163.12 and so these respecifications could be made in the first stage of our respecification search if they are deemed theoretical meaningful. However, all the modifications cannot be made simultaneously since a single respecification will have repercussions for the worth of others to the model as a whole. Even though the first two exceed our critical value, these respecifications cannot be made because they are not theoretically meaningful(for obvious reasons which have been stated earlier). So, we make the respecification using the third index ,that is, allowing MOTOR and COGNITION to covary. The results of adding such a path to our model has been discussed earlier.

Let's look at the next ten largest MIs after our first respecification search has been made.

Table 4.2: MIs for ORTHOPEDIC PATIENTS after first respecification

Modification	χ_2-value	P-value
COMP $\leftarrow$ EXPRESS	109.7483	0.000
EXPRESS $\leftrightarrow$ COMP	109.7474	0.000
EXPRESS $\leftrightarrow$ COMP	109.7464	0.000
BATH $\leftarrow$ TUBSH	83.1307	0.000
TUBSH $\leftrightarrow$ BATH	83.1303	0.000
TUBSH $\leftarrow$ BATH	83.1302	0.000
EAT $\leftarrow$ COG	49.1303	0.000
COG $\leftrightarrow$ EAT	49.1303	0.000
MOTOR $\leftrightarrow$ EAT	49.1299	0.000
BOWEL $\leftrightarrow$ BLADDER	39.0203	0.000

In the next stage of our respecification search, a modification will be considered if it exceeds a critical chi-square value with 134 (one fewer than the previous one) or 162.02 (assuming $\alpha = 0.05$). It can be seen that no modification index (as shown above) exceeds our value of 162.02 and so respecification ceases.

Below is the output for the STROKE PATIENTS (Refer to Appendix C for full model)

```
Model Chisquare =  534.51   Df =  135 Pr(>Chisq) = 6.2937e-49
 Goodness-of-fit index =  0.82674
 Adjusted goodness-of-fit index =  0.78053
 RMSEA index =  0.1019
 Bentler CFI =  0.88486
 SRMR =  0.20847
 AIC =  606.51
 BIC =  738.12
```

Let's look at the five (5) largest MIs and make the necessary modifications based on our criteria. It can be observed that, none of the modification indices above exceeds our critical

Table 4.3: MIs for STROKE PATIENTS

Modification	χ_2-value	P-value
COG $\leftarrow$ EAT	73.4597	0.000
COG $\leftrightarrow$ MOTOR	68.7774	0.000
COG $\leftarrow$ MOTOR	68.7774	0.000
MOTOR $\leftarrow$ COG	68.7774	0.000
MOTOR $\leftarrow$ INTERACT	67.3275	0.000

value with 135 degrees of freedom or 163.12 (assuming $\alpha = 0.05$). This means that no modification will be conducted on the STROKE data.

Chapter 5

Discussion

Generally, statistical significance testing in SEM hinges on the same principles as those of ANOVA, regression etc. The chi-square global test of statistical significance in SEM assesses whether our theoretical model fits our data, that is, whether our model-implied covariance matrix and observed covariance matrix are equal.

Hypothesis testing, however, differs from the traditional tests of statistical significance and this has repercussions for statistical power within the SEM framework. Mostly, in hypothesis testing, the null hypothesis depicts that there is no relationship (or no difference) between variables. Generally, we would want to reject the null and conclude that there are differences or (there exists a relationship) between variables. This logic is the reverse in SEM.

In SEM, we test the null hypothesis that our model-implied covariance matrix and the observed covariance matrix are equal. We would want to fail to reject the null in this scenario. We evaluate exact model fit by comparing the chi-square of the specified model to the critical value for the chi-square for its degrees of freedom. When the chi-square is statistically significant, we reject the null hypothesis that our model-implied covariance matrix and the observed matrix are equal.

The χ^2 is very sensitive to sample size. Due to its sensitivity to sample size, almost any model that has a large sample size will be rejected even if there is a minute amount of data misfit. Any SEM researcher wants to fail to reject the null hypothesis but having a large sample size always works against the researcher, that is, it provides more power to reject the null hypothesis that the model fits the data.

In our case, models for both Orthopedic and Stroke patients were rejected (that is, the

null hypothesis was rejected) even though they had different sample sizes. The respective sample sizes might be different but they are large and so the chi-square test of exact fit is likely to be rejected.

Our method for controlling type I error rate is quite logical but it could be overly conservative for large measurement models, as they often have hundreds of degrees of freedom and highly correlated indicators. However, it is sensible for many different SEMs and helps prevent overly capitalizing on chance when making modifications.

References

[1] Babakus, E. "The sensitivity of maximum likelihood factor analysis given violations of interval scale and multivariate normality," *Unpublished PhD dissertation* The University of Alabama.

[2] Bender, R and Lange, S. "Adjusting for multiple testing - when and how?" *Journal Of Clinical Epidemiology* 2001, Vol.54, pg. 343–349.

[3] Bollen, K.A and Noble, M.D. "Structural equation models and the quantification of behaviour," *Department of Sociology, University of North Carolina, Chapel Hill*

[4] Browne, M.W. "Covariance Structures," *Topics in Applied Multivariate Analysis, Douglas M. Hawkins, ed.* Cambridge, UK: Cambridge University Press, pg. 72–141.

[5] Cribbie A.R. "Evaluating the importance of individual parameters in structural equation modeling: The need for Type I error control," *Personality and Individual Differences*, 2000, Vol.29, pg. 567–577.

[6] Cribbie A.R. "Structural Equation Modelling," *A Multidisciplinary Journal*, 2007, pg. 98–112.

[7] Dunn, O.J. "Estimation of the means of dependent variables," *Annals of Mathematical Statistics*, 1958, Vol.29, pg. 1095–1111

[8] Games, P.A. "Multiple comparisons of means," *American Educational Research Journal*, 1971, Vol.8, pg. 531–565.

[9] Gerbing, D.W and Anderson, J.C. "The effects of sampling error and model characteristics on parameter estimation for maximum likelihood confirmatory factor analysis," *Multivariate Behavioral Research* 1985, Vol. 20, pg. 255–271.

[10] Gerbing, D.W and Anderson, J.C "Monte Carlo evaluations of goodness of fit indices for structural equation models," *Sociological Methods and Research*, 1992, Vol. 21, No. 2, pg. 132–160.

[11] Green, S.B., Thompson M.S and Poirer J. "An Adjusted Bonferroni Method for Elimination of Parameters in Specification Addition Searches , Structural Equation Modeling" *A Multidisciplinary Journal*, Psychology Press, London, 2009, pg. 18–39.

[12] Green S.B., Thompson M.S and Babyak M.A. "A Monte Carlo Investigation of Methods for Controlling Type I Errors with Specification Searches in Structural Equation Modeling" *Multivariate Behavioral Research*, 1998,Vol.33, No.3, pg. 365–383.

[13] Green S.B and Babyak M.A. "Control of Type I errors with Multiple Tests of Constraints in Structural Equation Modeling" *Multivariate Behavioral Research*, 1997,Vol.31, No.1, pg. 39–51.

[14] Hancock G.R. "A sequential Scheffe-type respecification procedure for controlling type I error in Exploratory Structural Equation Model Modification , Structural Equation Modeling" *A Multidisciplinary Journal*, 1999, Vol. 6, No.2, pg. 158–168.

[15] Hancock, G.R and Klockars, A.J. "The quest for α: developments in multiple comparison procedures in the quarter century since Games," *Review of Educational Research*, 1996, Vol. 66, pg. 269–306

[16] Hommel, G. and Bernhard, G. "Bonferroni procedures for logically related hypotheses," *Journal of Statistical Planning and Inference*, 1999, Vol. 82, pg. 119–128.

[17] Hu, L. and Bentler, P.M. "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," *Structural Equation Modeling*, 1999, Vol. 6, No. 1, pg. 1–55.

[18] Iacobucci D. "Structural Equation Modeling : Fit Indices, sample sizes, and advanced topic" *Journal Of Consumer Psychology*, 2009, Vol.20, pg. 90–98.

[19] Joreskog, K.G. "A general approach to confirmatory maximum likelihood factor analysis," *Psychometrika* 1969, Vol. 34, pg. 183–202.

[20] Joreskog, K.G. "Testing structural equation models. In K.A. Bollen and J.S. Long (Eds.), *Testing structural equation models*, 1993, pg. 294–316. Newbury Park, CA: Sage.

[21] Joreskog, K.G and Sorbom, D. "LISREL VI; Analysis of Linear structural relationships by Maximum Likelihood, Instrumental variables, and Least Squares," Uppsala: University of Uppsala.

[22] Kaplan, D. and Wenger, R.N "Asymptotic independence and separability in covariance structure models: Implications for specification error, power and model modification," *Multivariate Behavioral Research*, 1993, Vol. 28, pg. 467–482.

[23] Kirk, R.E "Experimental design: procedures for the behavioral sciences," 1995, Toronto, Canada: Brooks/Cole

[24] Kline, R.B. "*Principles and practice of structural equation modeling, 2nd ed.*," New York: Guildford

[25] Lord, F.M and Novick, M.R "Statistical Theories of Mental Tests" Reading: Addison-Wesley

[26] McArdle, J.J "Causal modeling applied to psychometric systems simulation," *Behavior Research Methods and Instrumentation*, 1980, Vol. 12, pg. 193–209.

[27] McArdle, J.J and McDonald, R.P "Some algebraic properties of the reticular action model," *British Journal of Mathematical and Statistical Psychology*, 1984, Vol.37, pg. 234–251.

[28] McCoach, D.S., Black, A.C and O'Connell, A. "Errors of Inference in Structural Equation Modeling," *Psychology in Schools*, 2007, Vol. 44, No.5, pg. 461–470.

[29] Muthen, B.O. "A General structural equation model with dichotomous, ordered categorical and continuous latent variable indicators," *Psychometrika* 1984, Vol. 49, pg. 115–132.

[30] Olsson, U. "Maximum likelihood estimation of the polychoric correlation coefficeint," *Psychometrika*, 1979, Vol. 44, No. 4, pg. 443–460.

[31] Olsson, U. "Maximum likelihood estimation of the Polychoric correlation coefficient," *Psychometrika*, 1979, Vol. 44,pg. 443–60.

[32] Pearson, K. and Pearson, E. "On Polychoric Coefficients of Correlation," *Biometrika*, 1922, Vol. 44, pg. 127–56.

[33] Ritchie-Scott, A. "On Polychoric Coefficients of Correlation," *Biometrika*, 1918, Vol. 18, pg. 342–53.

[34] Roy, S.N and Bose, R.C "Simultaneous confidence interval estimation," *Annals of Mathematical Statistics*, 1953, Vol. 24, pg. 513–536.

[35] Sidak, Z. "Rectangular confidence regions for the means of multivariate normal distributions," *Journal of the American Statistical Association* 1967, Vol. 62, pg. 626–633

[36] Sordom, D. "Model Modification," *Psychometrika*, 1989, Vol.54, No. 3, pg. 371–384.

[37] Spearman, C. "The proof and measurement of association between two thing," *The American Journal of Pychology*, 1987, Vol.100, pg. 441–471

[38] Stineman, M., Jette, A ., Fielder, R and Granger, C. "Impairment-specific dimensions within the functional independence measure," *Arch Phys Med Rehabil*, pg. 636–646

[39] Tukey, J.W "The problem of multiple comparisons," Unpublished manuscript, Princeton University, Department of Statistics.

[40] Ullman, J.B "Structural Equation Modeling (In : *Using Multivariate Statistics, Third edition*, B.G. Tabachnick and L.S. Fidell, Eds.)

Appendix A

Specifying the Model

```
> model.FIMSTROKE<-specifyModel()
1: COGSTROKE->COMPSTR,lam1,NA
2: COGSTROKE->EXPRESSSTR,lam2,NA
3: COGSTROKE->INTERACTSTR,lam3,NA
4: COGSTROKE->PSOLVINGSTR,lam4,NA
5: COGSTROKE->MEMORYSTR,lam5,NA
6: MOTORSTROKE->EATSTR,lam6,NA
7: MOTORSTROKE->GROOMSTR,lam7,NA
8: MOTORSTROKE->BATHSTR,lam8,NA
9: MOTORSTROKE->DRESSUPSTR,lam9,NA
10: MOTORSTROKE->DRESSLOSTR,lam10,NA
11: MOTORSTROKE->TOILSTR,lam11,NA
12: MOTORSTROKE->BLADDERSTR,lam12,NA
13: MOTORSTROKE->BOWELSTR,lam13,NA
14: MOTORSTROKE->BEDCHAIRSTR,lam14,NA
15: MOTORSTROKE->TOITRSFERSTR,lam15,NA
16: MOTORSTROKE->TUBSHOWERSTR,lam16,NA
17: MOTORSTROKE->WALKSTR,lam17,NA
18: MOTORSTROKE->STAIRSSTR,lam18,NA
19: COGSTROKE<->COGSTROKE,NA,1
20: MOTORSTROKE<->MOTORSTROKE,NA,1
21: COMPSTR<->COMPSTR,lam19,NA
```

```
22: EXPRESSSTR<->EXPRESSSTR,lam20,NA
23: INTERACTSTR<->INTERACTSTR,lam21,NA
24: PSOLVINGSTR<->PSOLVINGSTR,lam22,NA
25: MEMORYSTR<->MEMORYSTR,lam23,NA
26: EATSTR<->EATSTR,lam24,NA
27: GROOMSTR<->GROOMSTR,lam25,NA
28: BATHSTR<->BATHSTR,lam26,NA
29: DRESSUPSTR<->DRESSUPSTR,lam27,NA
30: DRESSLOSTR<->DRESSLOSTR,lam28,NA
31: TOILSTR<->TOILSTR,lam29,NA
32: BLADDERSTR<->BLADDERSTR,lam30,NA
33: BOWELSTR<->BOWELSTR,lam31,NA
34: BEDCHAIRSTR<->BEDCHAIRSTR,lam32,NA
35: TOITRSFERSTR<->TOITRSFERSTR,lam33,NA
36: TUBSHOWERSTR<->TUBSHOWERSTR,lam34,NA
37: WALKSTR<->WALKSTR,lam35,NA
38: STAIRSSTR<->STAIRSSTR,lam36,NA
39:
Read 38 records
>
```

Appendix B

ORTHOPEDIC OUTPUT

Parameter Estimates

	Estimate	Std Error	z value	Pr(>|z|)	
the1	0.91972	0.027056	33.9928	2.8449e-253	COMPOT <--- COGORTHO
the2	0.92430	0.026954	34.2913	1.0596e-257	EXPRESSOT <--- COGORTHO
the3	0.88249	0.027863	31.6730	3.6557e-220	INTERACTOT <--- COGORTHO
the4	0.89835	0.027524	32.6386	1.1624e-233	PSOLVINGOT <--- COGORTHO
the5	0.86329	0.028261	30.5473	6.1485e-205	MEMORYOT <--- COGORTHO
the6	0.42165	0.036412	11.5798	5.2180e-31	EATOT <--- MOTORORTHO
the7	0.52575	0.035355	14.8708	5.1016e-50	GROOMOT <--- MOTORORTHO
the8	0.61835	0.034199	18.0810	4.4955e-73	BATHOT <--- MOTORORTHO
the9	0.57776	0.034730	16.6355	3.8564e-62	DRESSUPOT <--- MOTORORTHO
the10	0.46753	0.035977	12.9952	1.3021e-38	DRESSLOOT <--- MOTORORTHO
the11	0.67976	0.033323	20.3994	1.6922e-92	TOILOT <--- MOTORORTHO
the12	0.57247	0.034797	16.4518	8.1430e-61	BLADDEROT <--- MOTORORTHO
the13	0.47113	0.035941	13.1083	2.9532e-39	BOWELOT <--- MOTORORTHO
the14	0.52378	0.035377	14.8056	1.3485e-49	BEDCHAIROT <--- MOTORORTHO
the15	0.68393	0.033260	20.5629	5.9052e-94	TOITRSFEROT <--- MOTORORTHO
the16	0.60092	0.034432	17.4525	3.2960e-68	TUBSHOWEROT <--- MOTORORTHO
the17	0.32760	0.037155	8.8169	1.1762e-18	WALKOT <--- MOTORORTHO
the18	0.25850	0.037577	6.8793	6.0132e-12	STAIRSOT <--- MOTORORTHO
the19	0.15412	0.010368	14.8640	5.6425e-50	COMPOT <--> COMPOT
the20	0.14567	0.010041	14.5080	1.0786e-47	EXPRESSOT <--> EXPRESSOT

```
the21 0.22121  0.013201  16.7576  4.9852e-63 INTERACTOT <--> INTERACTOT
the22 0.19296  0.011969  16.1214  1.8059e-58 PSOLVINGOT <--> PSOLVINGOT
the23 0.25473  0.014703  17.3256  3.0153e-67 MEMORYOT <--> MEMORYOT
the24 0.82221  0.042646  19.2800  7.9101e-83 EATOT <--> EATOT
the25 0.72359  0.038721  18.6872  6.2877e-78 GROOMOT <--> GROOMOT
the26 0.61764  0.034628  17.8364  3.6876e-71 BATHOT <--> BATHOT
the27 0.66620  0.036483  18.2603  1.7126e-74 DRESSUPOT <--> DRESSUPOT
the28 0.78141  0.041012  19.0532  6.1842e-81 DRESSLOOT <--> DRESSLOOT
the29 0.53792  0.031692  16.9733  1.2936e-64 TOILOT <--> TOILOT
the30 0.67228  0.036718  18.3091  7.0058e-75 BLADDEROT <--> BLADDEROT
the31 0.77804  0.040878  19.0333  9.0302e-81 BOWELOT <--> BOWELOT
the32 0.72566  0.038802  18.7014  4.8245e-78 BEDCHAIROT <--> BEDCHAIROT
the33 0.53225  0.031490  16.9021  4.3439e-64 TOITRSFEROT <--> TOITRSFEROT
the34 0.63889  0.035435  18.0300  1.1338e-72 TUBSHOWEROT <--> TUBSHOWEROT
the35 0.89268  0.045493  19.6223  9.9739e-86 WALKOT <--> WALKOT
the36 0.93318  0.047141  19.7953  3.2650e-87 STAIRSOT <--> STAIRSOT

 Iterations =  51
```

Appendix C

STROKE OUTPUT

Parameter Estimates

	Estimate	Std Error	z value	Pr(>|z|)	
lam1	0.91376	0.045639	20.0213	3.5927e-89	COMPSTR <--- COGSTROKE
lam2	0.91289	0.045672	19.9881	6.9886e-89	EXPRESSSTR <--- COGSTROKE
lam3	0.90357	0.046014	19.6367	7.5069e-86	INTERACTSTR <--- COGSTROKE
lam4	0.87780	0.046939	18.7011	4.8466e-78	PSOLVINGSTR <--- COGSTROKE
lam5	0.91125	0.045732	19.9258	2.4319e-88	MEMORYSTR <--- COGSTROKE
lam6	0.61972	0.055102	11.2467	2.4034e-29	EATSTR <--- MOTORSTROKE
lam7	0.69372	0.053307	13.0138	1.0208e-38	GROOMSTR <--- MOTORSTROKE
lam8	0.75123	0.051729	14.5224	8.7366e-48	BATHSTR <--- MOTORSTROKE
lam9	0.71961	0.052617	13.6763	1.4058e-42	DRESSUPSTR <--- MOTORSTROKE
lam10	0.68499	0.053532	12.7958	1.7309e-37	DRESSLOSTR <--- MOTORSTROKE
lam11	0.81661	0.049731	16.4205	1.3639e-60	TOILSTR <--- MOTORSTROKE
lam12	0.67393	0.053812	12.5236	5.5428e-36	BLADDERSTR <--- MOTORSTROKE
lam13	0.55727	0.056425	9.8762	5.2784e-23	BOWELSTR <--- MOTORSTROKE
lam14	0.71462	0.052753	13.5466	8.2964e-42	BEDCHAIRSTR <--- MOTORSTROKE
lam15	0.72510	0.052466	13.8202	1.9249e-43	TOITRSFERSTR <--- MOTORSTROKE
lam16	0.68393	0.053559	12.7695	2.4259e-37	TUBSHOWERSTR <--- MOTORSTROKE
lam17	0.49991	0.057495	8.6949	3.4704e-18	WALKSTR <--- MOTORSTROKE
lam18	0.44499	0.058395	7.6203	2.5313e-14	STAIRSSTR <--- MOTORSTROKE
lam19	0.16504	0.017929	9.2054	3.4037e-20	COMPSTR <--> COMPSTR
lam20	0.16663	0.018040	9.2371	2.5327e-20	EXPRESSSTR <--> EXPRESSSTR

```
lam21 0.18355  0.019238   9.5412 1.4122e-21 INTERACTSTR <--> INTERACTSTR
lam22 0.22946  0.022631  10.1393 3.6980e-24 PSOLVINGSTR <--> PSOLVINGSTR
lam23 0.16962  0.018248   9.2951 1.4700e-20 MEMORYSTR <--> MEMORYSTR
lam24 0.61595  0.054512  11.2992 1.3236e-29 EATSTR <--> EATSTR
lam25 0.51875  0.047216  10.9865 4.4365e-28 GROOMSTR <--> GROOMSTR
lam26 0.43565  0.041070  10.6076 2.7469e-26 BATHSTR <--> BATHSTR
lam27 0.48216  0.044497  10.8359 2.3264e-27 DRESSUPSTR <--> DRESSUPSTR
lam28 0.53079  0.048116  11.0316 2.6916e-28 DRESSLOSTR <--> DRESSLOSTR
lam29 0.33315  0.033721   9.8795 5.1074e-23 TOILSTR <--> TOILSTR
lam30 0.54582  0.049240  11.0849 1.4848e-28 BLADDERSTR <--> BLADDERSTR
lam31 0.68945  0.060074  11.4768 1.7261e-30 BOWELSTR <--> BOWELSTR
lam32 0.48932  0.045027  10.8672 1.6526e-27 BEDCHAIRSTR <--> BEDCHAIRSTR
lam33 0.47423  0.043910  10.8002 3.4358e-27 TOITRSFERSTR <--> TOITRSFERSTR
lam34 0.53224  0.048224  11.0368 2.5380e-28 TUBSHOWERSTR <--> TUBSHOWERSTR
lam35 0.75009  0.064680  11.5969 4.2752e-31 WALKSTR <--> WALKSTR
lam36 0.80198  0.068633  11.6851 1.5186e-31 STAIRSSTR <--> STAIRSSTR

 Iterations =  51


> modIndices(sem.FIMST)

 5 largest modification indices, A matrix:
      COGSTROKE<-EATSTR   COGSTROKE<-MOTORSTROKE   MOTORSTROKE<-COGSTROKE
               73.45967                 68.77739                 68.77739
MOTORSTROKE<-INTERACTSTR  MOTORSTROKE<-EXPRESSSTR
               67.32746                 63.57239
```

```
5 largest modification indices, P matrix:
  MOTORSTROKE<->COGSTROKE     DRESSLOSTR<->DRESSUPSTR
                 68.77739                    42.43594
   TUBSHOWERSTR<->BATHSTR TUBSHOWERSTR<->TOITRSFERSTR
                 42.07205                    33.83882
       COGSTROKE<->EATSTR
                 22.59519
```

Appendix D

Code

```
library(polycor)
hetcor(FIM,ML=FALSE)
hcor=function(data) hetcor(data,std.error=FALSE)$correlations
r.FIM=hcor(FIM)

sem.FIMST<-sem(model.FIMSTROKE,r.FIM,N=286,method=‘‘GLS’’)
```

Curriculum Vitae

John Appiah Kubi was born on April 25, 1984. The first son of Kwaku Appiah Kubi and Agatha Quayson, he graduated from St. John's High School, Sekondi, Ghana, in the Spring of 2002. He entered Kwame Nkrumah University Of Science and Technology, (KNUST), Kumasi, Ghana in the Fall of 2003 and obtained his Bachelor's degree in Mathematics. In 2007, he was hired as a teaching assistant in the Department of Mathematics, KNUST, Kumasi, Ghana as a national service personnel. In the Fall of 2008, after honorably discharging his duties as a service personnel, John resumed his studies at KNUST, Kumasi, Ghana as a master's student. While pursuing his master's degree in Mathematics, he worked as a Graduate Teaching Assistant. He received his master's degree in Mathematics in the Spring of 2010.

In the Fall of 2011, John moved to El Paso and entered the Graduate School of The University of Texas at El Paso. While pursuing a master's degree in Statistics, he worked as a Graduate Teaching Assistant in Mathematical Science deparment. He is willing to start his PhD in Biostatistics in the Fall of 2013.

Permanent address: P.O.Box KJ 64
Kojokrom, Sekondi, Ghana.