1-1-2016

Estimating Effects Of Non-Normality In Assessing Structural Equation Model Fit For Use Of Physical Science Data

Sarah Rose
Wayne State University,

ESTIMATING EFFECTS OF NON-NORMALITY IN ASSESSING STRUCTURAL EQUATION MODEL FIT FOR USE OF PHYSICAL SCIENCE DATA

by

SARAH ALTA ROSE

DISSERTATION

Submitted to the Graduate School

of Wayne State University

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2016

MAJOR: EDUCATION (Evaluation and Research)

Approved By:

____
Advisor Date

____

____

____

____

DEDICATION

I would like to dedicate this dissertation to my family who supported and encouraged me through this, as in every endeavor. My grandmother, Fannie Rose, passed away shortly before the completion of this work; her love and support are truly missed. My grandmother, Marian Frankel, loved education and inspired me to always aim higher academically. My grandfathers, Sam Rose and Leonard Frankel, always gave me their love and support. My father, Mitchell Rose, never ceased to remind me that a formal education doesn't stop until after the doctorate. My mother never ceased to remind me that completing my dissertation was very important. My stepfather has always been an inspiration for me in the continuation of my academic career. My sister, Chaya Thav, reminded me to hold my breath during the difficult times, and that school would be done before I knew it. To my stepmother, Svetlana Rose, and the rest of my brothers and sisters.

ACKNOWLEDGMENTS

I am very grateful to my major advisor, Dr. Barry Markman, who led me through this dissertation process. Thank you for all your help and support. I am also grateful to the other members of my committee, Drs. Monte Piliawsky, Boris Shulkin, Jamie Gleason, and Kraig Warnemuende, who had an important role in the dissertation process. I would like to thank Drs. Shlomo Sawilowsky, Robert Partridge, Bruce Zumbo, and Ryoungsun Park for their consultation and advice concerning the results of the Monte Carlo analyses.

A special thanks to Dr. Shlomo Sawilowsky who had an important role in my educational and professional development.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Background of the Study

Structural Equation Modeling (SEM) is a relatively new statistical methodology that is beginning to be established in the professional field of statistics. Its foundational theory was published by Wright (1918), where a path analysis was used to model the bone size of rabbits. However, the novelty of the methodology was such that SEM was not accepted by researchers until the 1960s or 1970s (Matsueda, 2011). This coincided with increasing use of computers, allowing for the more practical use of the complicated matrix models by standard researchers.

The development of more complicated analytical procedures was inevitable. Hoyle (1995) indicated, “with the increasing complexity and specificity of research questions in the social and behavioral sciences…has come increasing interest in SEM as a standard approach to testing research hypotheses” (p. 1). Indeed, with the complex nature of many modern research models, it is imperative to use a data analysis tool that allows the most flexibility in the analysis to confirm a best interpretation of the model results. SEM is a powerful tool that can be used to explore data for the purpose of improving the understanding of the interactions, reliability, and general characteristics. It allows for a more complete and comprehensive analysis compared to other research methodologies (such as multiple regression) because it allows freedom in the evaluation of several model construct relationships simultaneously (Alavifar, 2012). This advantage should not be underestimated. The ability to take 5, 10, 20, or 100 variables and analyze them together using one test without the necessity for Bonferonni or similar corrections allows for considerable increase in statistical power.

SEM has the unique capability to model relationships between variables and to estimate error. SEM can therefore be considered as rather a “union” (p. 3) between path analysis and factor analysis (Gullen, 2000). Modelling error in SEM is a unique advantage. By virtue that error is explicit in the SEM model, as opposed to the implicit implication of error via other methodologies, using SEM can result in a more realistic, reliable, and comprehensive model of the data.

SEM models are developed by determining relationships between observed and/or latent variables to develop an initial model. The model is analyzed to determine whether it is an appropriate approximation of the data construct. If the model is concluded to be an appropriate approximation, it is analyzed to ascertain the magnitude and direction of relationships between the different variables.

Kline (2011) set forth six steps to developing a SEM model. They are (1) specifying the model, (2) evaluating the model identification, (3) selecting the measures and collecting and screening the data, (4) estimating the model, (5) re-specifying the model, and (6) reporting the results.

Probably, the most essential step from the six mentioned above is Step 4. This includes assessing the model to determine how well it represents the data. Many SEM models can be developed that represent the data to a degree; however, a good model will be the best fit representation. To this end, model fit statistics were developed. These statistics result in a quantitative analysis of model fit that allows researchers to determine how well the model fits the data in an objective manner.

The matter of how to develop these fit statistics and which are the best to use has been a topic of great discussion. Kline (2011) indicated that “For at least 30 years the literature has carried an ongoing discussion about the best ways to test hypotheses and assess model fit” (p. 190). There are dozens of fit indices, and each one is a measure of appropriate model fit to the data. For example, IBM’s SPSS Amos Graphics (version 22) provides 20 different fit indices; however, there are dozens of different fit indices that can be considered (Kline, 2011). Most researchers agree that the Chi-Squared test (or Cmin as indicated Amos Graphics) is a basic evaluation of model fit for SEM (Kline, 2011 and Hoyle, 1995) and should be evaluated first. If this test indicates a bad fit, it should weigh considerably on the researcher’s assessment of the model fit.

Other common fit indices include the Root Mean Square Error of Approximation (RMSEA), the Standardized Root Mean Square Residual (SRMR), and the Comparative Fit Index (CFI). These are common fit indices that were recommended by Kline (2011), Hoyle (1995), Byrne (1994), and Hooper, Coughlan, and Mullen (2008). These fit indices are provided by Amos Graphics and are therefore easily obtained. They are discussed further in Chapter 2.

However, other fit indices do exist. These include the Goodness-of-Fit Index (GFI), Adjusted Goodness of Fit Statistic (AGFI), Root Mean Square Residual (RMR), and others (Hooper, Coughlan, & Mullen, 2008).

Each fit index is unique and measures model fit in different manners. For example, the Chi-Square test is based on the “magnitude of discrepancy” (p. 53) between the expected data and the actual data (Hooper, Coughlan, & Mullen 2008).

This test is based on the overall model fit, as opposed to the incremental fit (as will be discussed below). Although the Chi-Square test can be used to assess any model fit distribution; in SEM, the Chi-Square test generally is used to determine variance of the data from normality. The Chi-Squared test has several limitations that affect the Chi-Square values and can provide erroneous approximation of fit. Factors that can inflate or deflate the Chi-Square values include high correlation among observed variables, unique variances among variables, and large samples sizes (Kline, 2011). Additionally, the Chi-Square test gives little information as to the extent that model does not fit (Byrne, 1994). As such, additional statistical measures are necessary to determine model fit approximation of the SEM.

The fit indices for SEM have different limitations and boundary conditions. The necessity for numerous fit indices can be explained in two ways. Firstly, fit indices are greatly important in the performance of any SEM. SEM that is an improper fit to the data would provide inaccurate or erroneous results, and possibly indicate relationships that do not exist. Secondly, the process to performing SEM, the complexity of the variable matrixes and the sheer volume of analysis required, indicate a necessity for numerous fit index models. As the process is rigorous and complicated, so too the fit indexes are difficult to simplify. There is currently no single fit index that encompasses all the different indices in one comprehensive test. (Gullen, 2000)

The complexity of analyzing the fit indices and the plethora of index tests from which to form a model fit assumption, make it necessary to determine when models are truly a good fit to the data.

Hooper, Coughlan, and Mullen (2008) indicated:

> Given the plethora of fit indices, it becomes a temptation to choose those fit indices that indicate the best fit…This should be avoided at all costs as it is essentially sweeping important information under the carpet. (p. 56)

Hooper, Coughlan, and Mullen (2008) recommended several common fit indices to be considered, among which are listed the CFI, the RMSEA and the Chi-Square tests as the most common. However, the tests listed above were developed primarily for social and behavioral science research where the baseline assumption for distribution is normal. It is questionable whether these variables provide a good indication of model fit under different distributions that are common in the physical sciences (i.e. exponential, logarithmic, or uniform).

Problem Statement

The purpose of this study is to evaluate the sensitivity of selected fit index statistics in determining model fit when the distribution varies from normality, as is typically true of data research for the physical sciences. SEM is already being applied to many physical science research problems and the reliability and power of the model fit indices is questionable. Gullen (2000) discussed how "Non-normal data pose problems in structural equation models even if the data are continuous" (p. 19). Gullen (2000), however, did not consider the extent to which the problem is imposed, or which distributions perform better than others. The extent and power of the fit indices in estimating the SEM model fit when normality is violated is therefore of interest.

Theoretical mathematical and physical science distributions and applied physical science data sets will be obtained. Tests for normality will be conducted on the applied data sets. These data sets will be sufficiently large to serve as proxies for typical physical science distributions. The theoretical distributions and applied data sets will then be randomly sampled and subjected to RMSEA, SRMR, and CFI model fit index tests.

Assumption

1. Data sets from real variables and hypothetical distributions will be sampled. It is assumed they are representative of common conditions.

Limitations

1. There are dozens of different fit indices; however, only three common fit indices (RMSEA, SRMR, and CFI) will be used in this study.
2. There are limitless distributions possible. In this study, only three or four variables encompassing different distributions from several real and hypothesized data sets will be analyzed.
3. The results of the fit indices will vary based on the sample size and whether the sample size is balanced or unbalanced between variables. This study will be limited to three or four sample size groups and to balanced sample sizes between variables.
4. The results of the fit indices will vary based on the chosen alpha level. In this study, only one alpha level (of 0.05) will be used.

Definition of Terms

1. Observed Variables: Variables consisting of data that are measurable and provided directly from an instrument.
2. Latent Variables: Variables that consist of a combination of several observed variables, similar to a construct.
3. Model Specification: The process of arranging the observed and latent variables into an SEM model, and specifying their relationships (correlations, errors, and direction).
4. Model Identification: The process of identifying the degrees of freedom of an SEM model.
5. Correlation: The relationship or association between two variables. The correlation can be quantified using the correlation coefficient.
6. Error: Deviation of the data from expected value that cannot be attributed to a model or any substantial explanation.

CHAPTER 2 REVIEW OF LITERATURE

Overview of SEM

SEM is a comprehensive data analysis technique that allows researchers greater flexibility in determining the magnitude and direction of relationships between variables. A primary advantage of SEM lies in the creation of a model. These models can graphically provide quantitative values of variable relationships, correlations, and even error. Other advantages were offered by Chin (1998). Benefits of SEM over "first-generation techniques" (p. vii) (such as factor analysis, principal components analysis, multiple regression, and discriminant analysis) are listed below:

1. Ability to model relationships between variables
2. Ability to model error in observed variables
3. Ability to conduct a priori tests against empirical data (such as in Confirmatory Factor Analysis)
4. Hesketh, Skrondal, & Pickles (2004) offered a fourth advantage in that SEM can be used in the development of latent variables of "hypothetical construct" (p. 168).

SEM can be considered as a conglomeration of several common "first-generation" (Chin, 1998, p. vii) statistical approaches. In running SEM, statistical approaches that are simulated include correlation, multivariate regression, path analysis, maximum likelihood, generalized least squares, and factor analysis. It is for this reason that the SEM methodology contains many similarities with the conventional analytical procedures.

These similarities include:

1. Both are general linear models (Kline, 2011).
2. Analyses are only valid if boundary conditions are met (Weston & Gore, 2006)
3. The techniques do not imply causality (Weston & Gore, 2006)
4. Researchers can misuse the SEM just as with other, more classical, analytical procedures (Weston & Gore, 2006).

Weston and Gore (2006) paradoxically stated:

> "Just as researchers are free (although not encouraged) to conduct several different multiple regression models until they find a model to their liking, they can also analyze models in SEM, identify and remove weaknesses in the model, and then present the revised model as if it were the originally hypothesized model." (p. 723)

One common way that researchers can misuse models includes publicizing the fit index measures that indicate a good model fit and neglecting the indices that indicate a poor fit. As indicated above, there are a plethora of model fit indices and each one measures model fit in a different way. Additionally, as the distribution of the samples vary from the boundary conditions, as set forth by the SEM computer software (i.e. normality); the fit indices can result in erroneous assessments that can inflate model fit statistics. It is therefore of importance to understand how variations from normality affect the model fit equations.

Kline's (2011) Six Steps to Performing SEM

SEM is a relatively new statistical procedure that is tailored for a rigorous analytical approach of data research. Even using a computer program to solve for the matrix algebra and determine the output results, the approach for analyzing the results and for selecting the measures is complex. Kline (2011) simplified the procedures by identifying six steps to performing SEM. These steps, stated in Chapter 1, are explained in fuller detail below.

<u>Step 1 - Specifying the Model:</u> Specifying the model includes analysis of the relationships between the variables and drawing a model diagram. Model diagrams include endogenous and exogenous variables that are typically represented by arrows, indicating direction. Correlations are typically represented by a circular, two-direction arrow, indicating a strong relationship between two variables. SEM models assume that variables without correlations are not highly correlated. Ideal SEM models include primary variables that are moderately correlated with coefficient values ranging between 0.4 and 0.7.

<u>Step 2 - Model Identification:</u> Model identification is necessary prior to beginning model estimation to determine whether it is "*theoretically* possible for the computer to derive a unique estimate of every model parameter." (Kline, 2011, p. 93). An under-identified model cannot provide reliable model results. Therefore, prior to model estimation, it is imperative to determine that the model is over-identified or just-identified. Gullen (2000) compared this to algebra, as numerous equations and variables are required. "Unlike in algebra, however, there is a benefit to having more equations than variables. These over-identified models permit the calculation of fit

statistics for the evaluation of model fit." (p. 1). In other words, just-identified models can allow for a unique solution but the fit indices would not provide valid results for determination of model fit (see Step 4).

Under-identified, just-identified, and over-identified designations are a function of the model complexity. SEM analyses generally provide improved results with a more parsimonious model. A parsimonious model is a model with a larger degree of freedom. The degrees of freedom can be calculated per Kline (2011) using the following equation:

$$df_m = p - q \qquad \text{Eq. (1)}$$

$$p = v\left(v+1\right)/2 \ , \qquad \text{Eq. (2)}$$

where v = number of observed variables, and q = number of estimated parameters.

A model with fewer estimated parameters (i.e. correlations, error terms, etc.) and many observed variables would be more parsimonious. Parsimony is preferable, and many model fit indices include an adjustment to account for parsimonious models (refer to the subsection "Approximate Fit Indices" below).

Step 3 – Selecting the Measures: Selecting the measures and collecting/preparing data involve determining whether the data are appropriate for an SEM model. This includes ensuring that extreme collinearity does not occur. Extreme collinearity occurs when the primarily (latent) variables are highly correlated with each other. This step also includes dealing with outliers and missing data as prescribed by standard statistical procedures.

Step 4 – Model Estimation: Model estimation means determining whether the SEM is a good estimation of the data. This involves determining model fit based on appropriate fit indices, interpreting parameter estimates, and considering equivalent or near-equivalent models. Fit indices are tests that indicate whether the model is statistically a good fit for the data. There are dozens of different fit indices. Indeed, new ones are being developed with regularity (Kline, 2011). Determining appropriate model fit based on fit indices can be a complicated procedure. This will be further discussed below.

Interpreting model parameters includes examination to determine whether the model properly represents the data via a low coefficient of non-determination and high standardized regression weights. If the model is a poor or basically adequate fit, it is appropriate to look at several similar models to determine whether the new models are better.

Step 5 – Re-specifying the Model: Re-specifying the model includes changing the SEM to determine whether the re-specified model is a better fit for the data. This can include either dropping an arrow or dropping/adding a correlation.

Step 6 – Reporting the Results: Reporting the results includes publishing or writing a report of the results of the research and specifying a correlation or covariance matrix so that the results can be reproduced by another researcher, if desirable.

The purpose of the six steps above is to ensure that a model contains appropriate variable measures and variable relationships and provides the best model representation of the data. A SEM is only as good as the data fit and error magnitude. A poor model can result in false determinations of variable relationships and causal direction.

SEM and the Social Behavioral Sciences

SEM was specifically developed for the use of analysis of social and behavioral science data. Weston and Gore (2006) discussed how the necessity for developing SEM methodology was born on the complex and multivariable nature of social and behavioral science research.

> Weston and Gore (2006) wrote:
>
> “For example, noting the mostly univariate nature of extant eating-disorder research, Tylka and Suich (2004) hypothesized that eating-disorder patterns in adult women were a function of personal, socio-cultural, and relational factors. Furthermore, they offered hypotheses about how these factors interact in complex ways to explain symptom severity” (p. 719)

Additional examples of the complex nature of the variables and their relationships were discussed. Weston and Gore (2006) indicated researchers were reluctant to perform studies on complex issues because multivariate analysis was simply insufficient to meet their demands.

The boundary conditions for performing SEM and determining model fit are steeped in the conditions typical of social and behavioral sciences. These boundary conditions include multivariate normality (Gullen, 2000; Kline, 2011; Reinartz, Echambadi, & Chin, 2002; Tomarken & Waller, 2005). Although standard for most social and behavioral science parametric statistical tests, SEM modeling does not require homogeneity of variance (Baozzi & Youjae, 1989).

SEM and the Physical Sciences

Due to the capability of improving quality of life by analyzing data for complex research studies, SEM is increasingly being used in physical science research (Kelly, 2011 and Ewing, Hamidi, Gallivan, Nelson, & Grace, 2014). There are many similarities between social and behavioral science and physical science data that make this transfer of methodologies apparently appropriate. Both data sets are parametric, can be assigned descriptive statistic values, can be formulated to provide frequency diagrams, and can be used with nonparametric tests.

However, physical science data differ from the social behavioral science in several ways. Construct validity does not exist in the physical sciences, as the science deals with physical data that can be measured. Therefore, there is no need for the development of a measure for a hypothetical construct. Furthermore, physical science data have different distributions than that of social and behavioral science. Although social and behavioral science data tend to be centered on the normal distribution, physical science data can have virtually almost any distribution. The following graphs were taken from several articles, and show various distributions for physical science research.

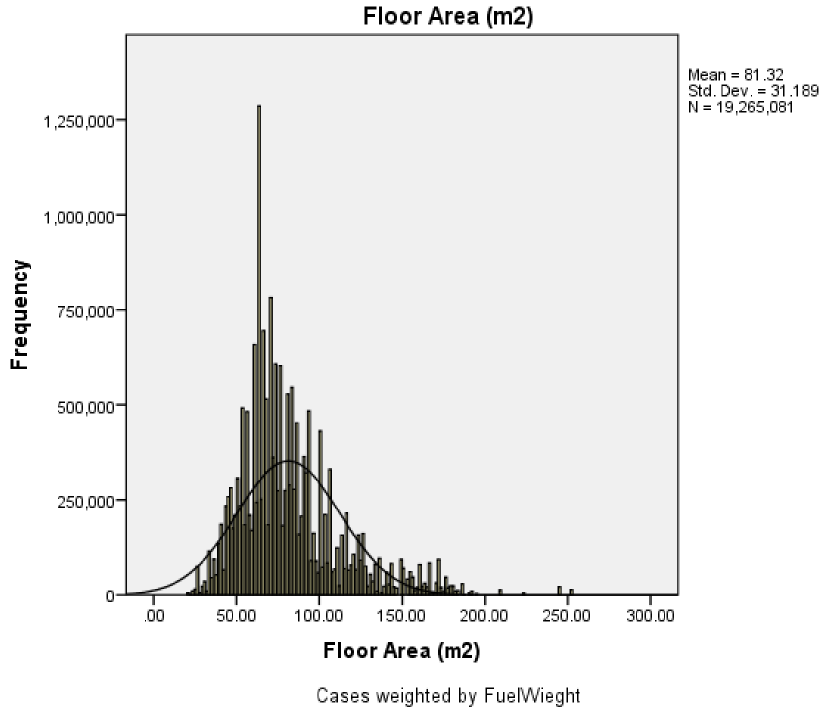

Figure 1. Ground Floor Area Frequency [Data Provided By Kelly (2011)].

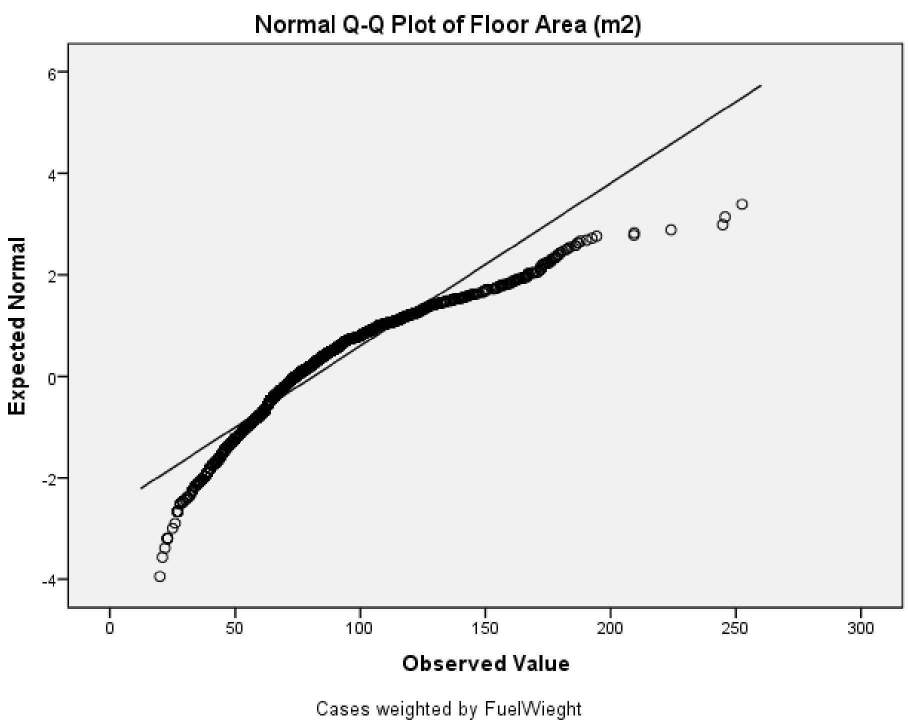

Figure 2. Ground Floor Area Q-Q Plot [Data Provided By Kelly (2011)].

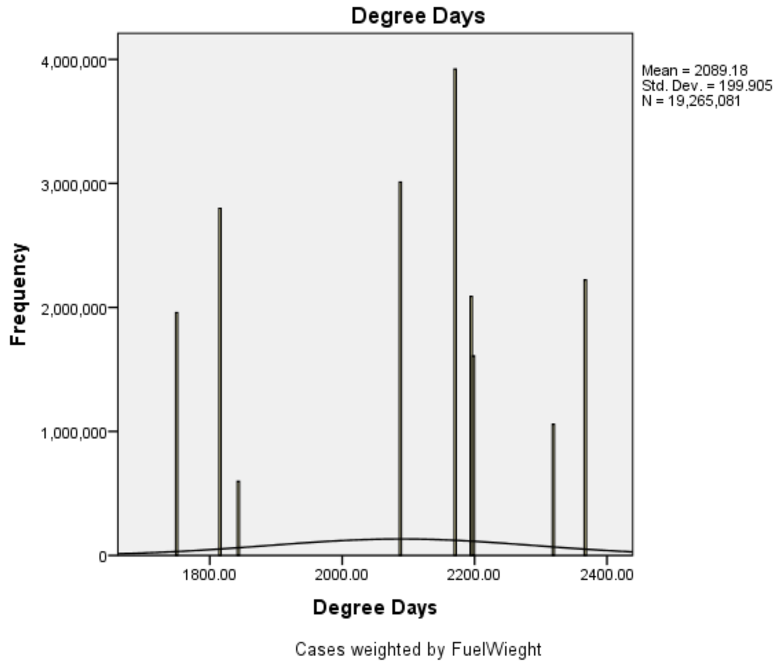

Figure 3. Degree Day Frequency [Data Provided By Kelly (2011)].

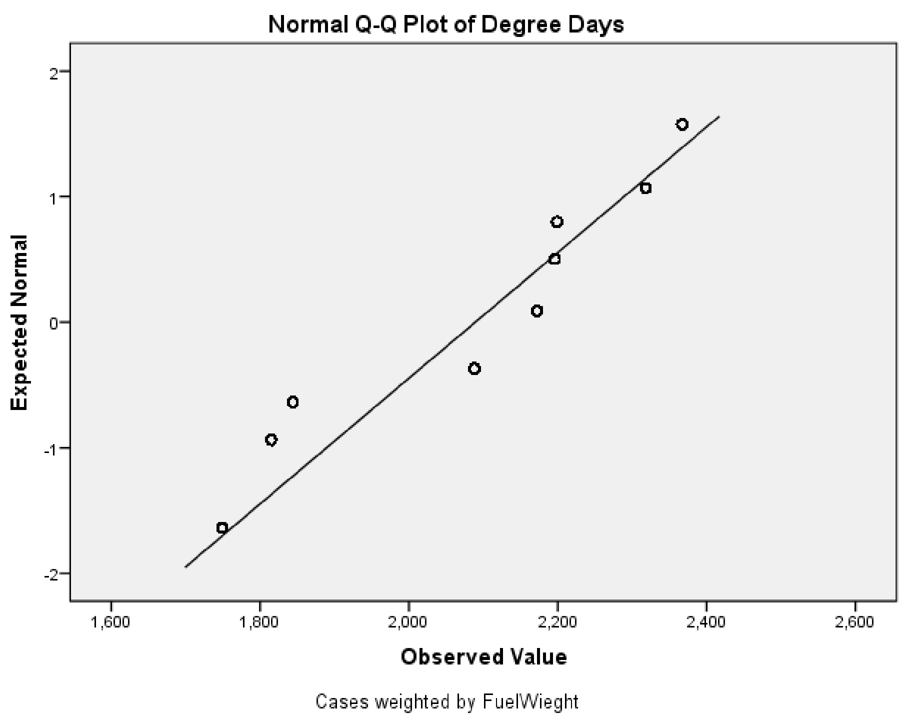

Figure 4. Degree Day Q-Q Plot [Data Provided By Kelly (2011)].

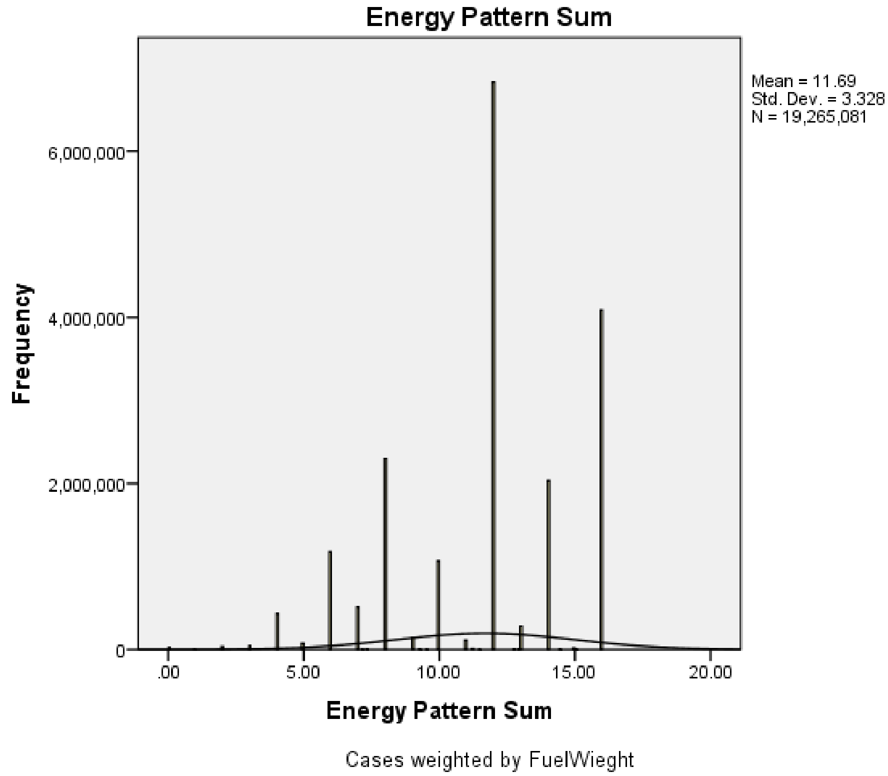

Figure 5. Energy Pattern Frequency [Data Provided By Kelly (2011)].

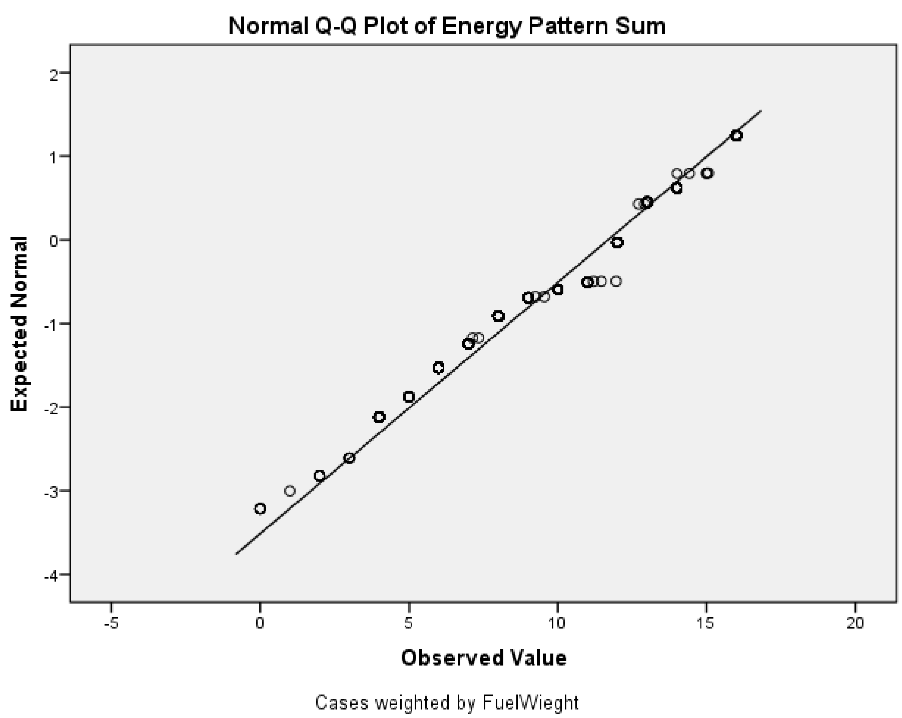

Figure 6. Energy Pattern Q-Q Plot [Data Provided By Kelly (2011)].

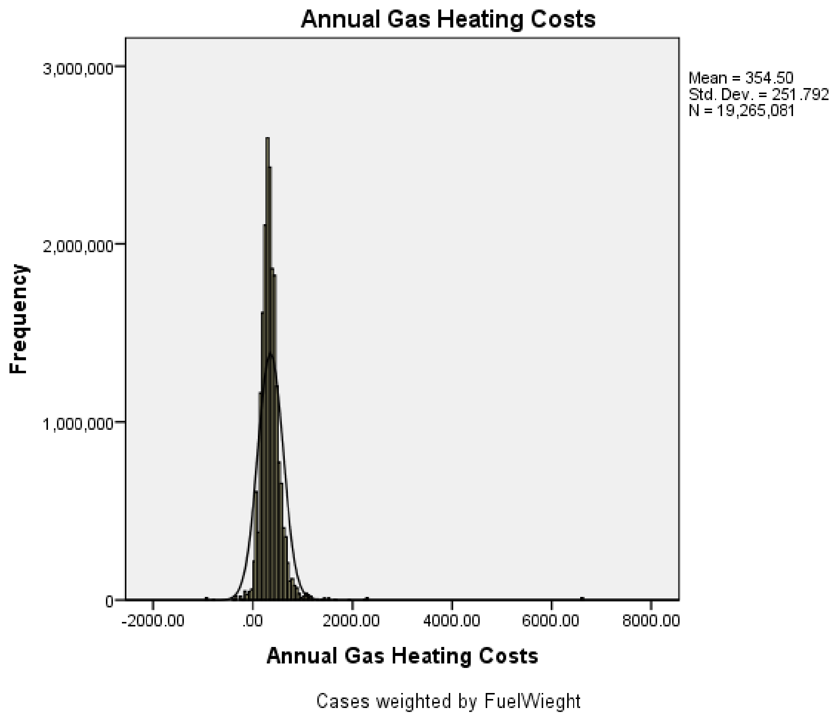

Figure 7. Total Annual Cost Frequency [Data Provided By Kelly (2011)].

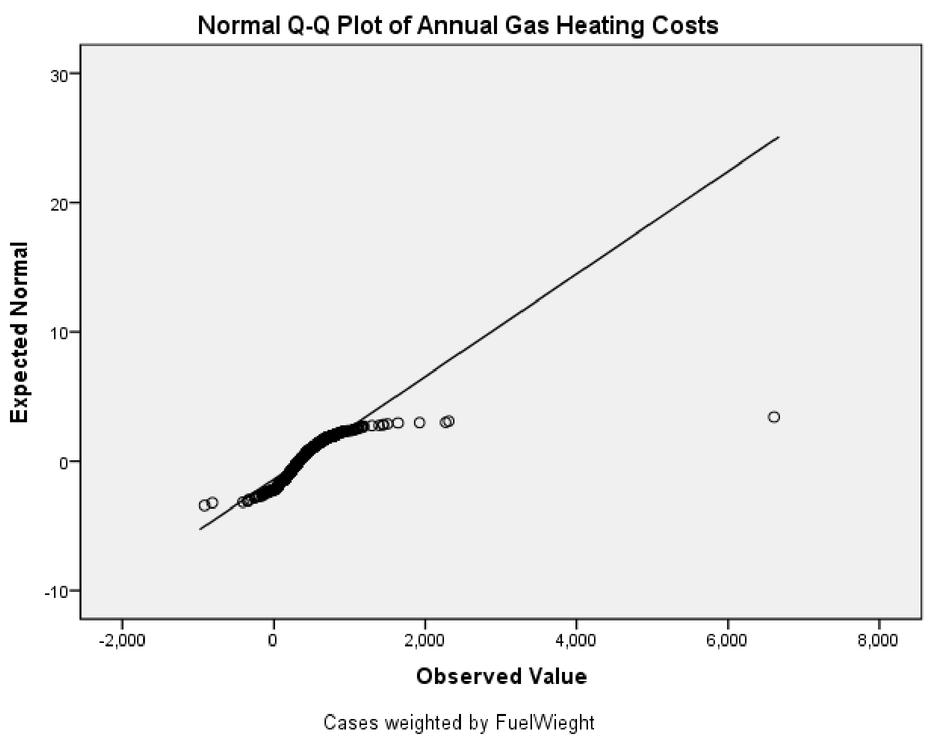

Figure 8. Total Annual Cost Q-Q Plot [Data Provided By Kelly (2011)].

Table 1

Test of Normality for Variables from Figures 1, 3, 5, and 7

	Kolmogorov-Smirnov[a]		
	Statistic	df	Sig.
Floor Area (m2)	.112	19265081	.000
Degree Days	.226	19265081	.000
Energy Pattern Sum	.226	19265081	.000
Annual Gas Heating Costs	.132	19265081	.000

a. Lilliefors Significance Correction

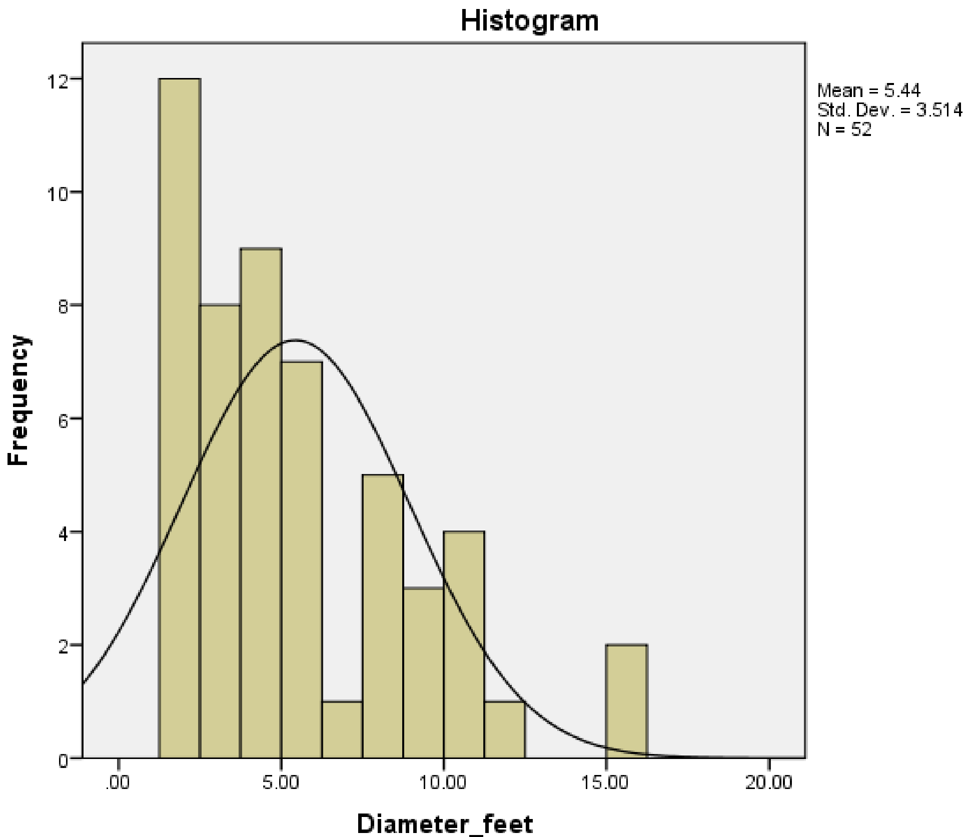

Figure 9. Diameter of Sewer Frequency
[Data Provided By Sinfield & Einstein (1998)].

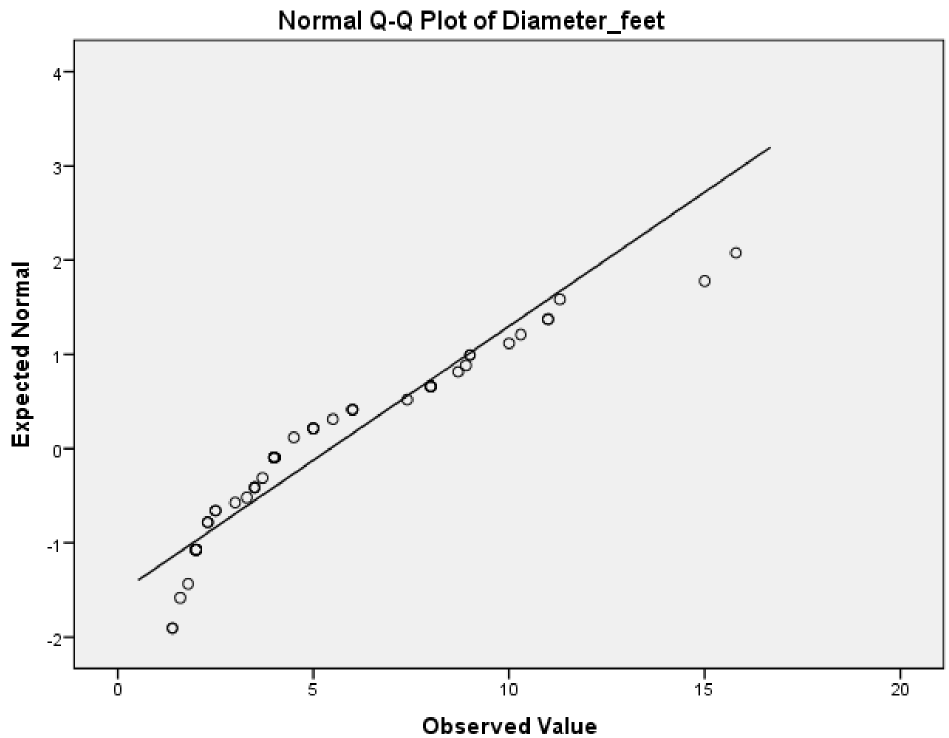

Figure 10. Diameter of Sewer Q-Q Plot
[Data Provided By Sinfield & Einstein (1998)].

Table 2

Test of Normality for Variables from Figure 8

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Diameter_feet	.197	52	.000	.886	52	.000

a. Lilliefors Significance Correction

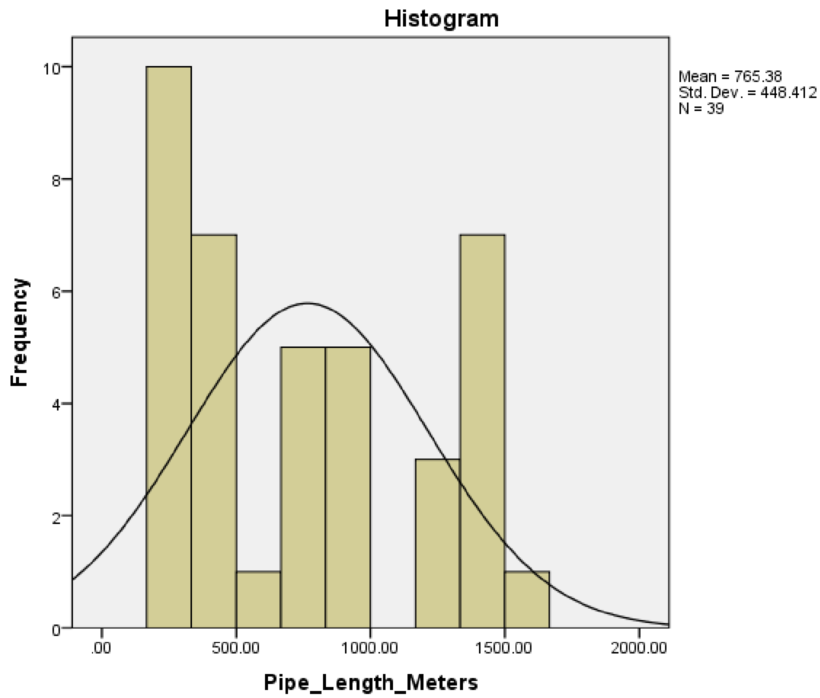

Figure 11. Tunnel Pipe Length Frequency
[Data Provided By Yeh, Lin, & Tsai (2008)].

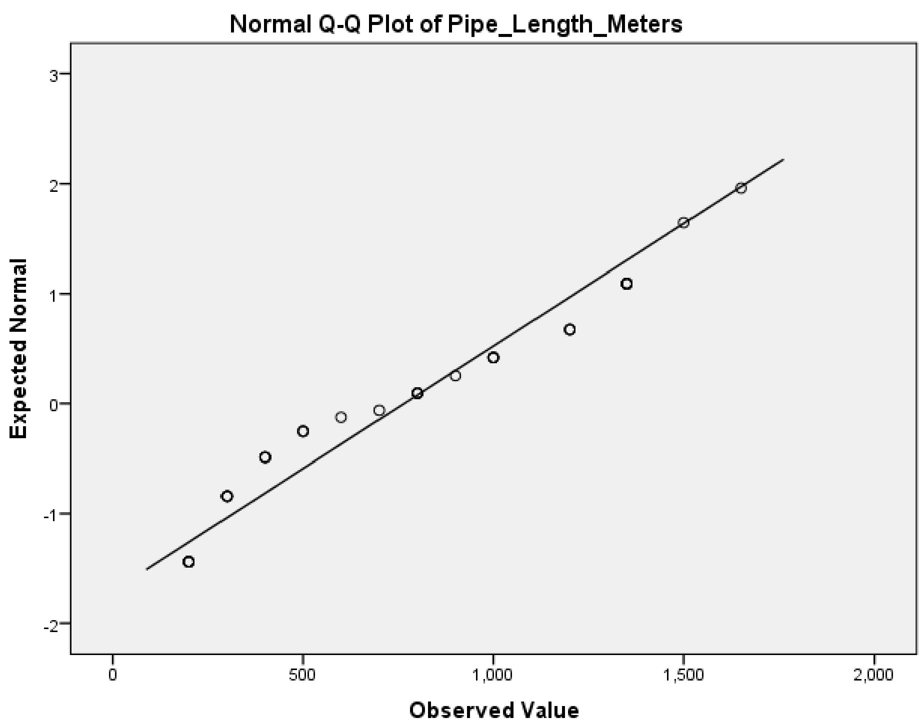

Figure 12. Tunnel Pipe Length Q-Q Plot
[Data Provided By Yeh, Lin, & Tsai (2008)].

Table 3

Test of Normality for Variables from Figure 10

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pipe_Length_Meters	.159	39	.014	.910	39	.004

a. Lilliefors Significance Correction

As can be seen, the distributions above are non-normal. In Tables 1, 2, and 3, the tests of normality (Kolmogorov-Smirnov and Shapiro Wilk) significance values are less than the alpha value of 0.05. Therefore, the null hypothesis that the data are a normal distribution is rejected and the data are considered to be non-normal (to a statistically significant degree).

It becomes apparent that data from the physical sciences are typically non-normal. Hence, the question arises of how well would SEM perform using physical science data that are non-normal? Part of determining the answer to that question would include examination of the fit indices that measure the model fit of the data. These fit indices are essential in estimating the SEM model, for without a proper model fit, the SEM model would become ineffectual, regardless of the other output results.

Fit Indices Background

Model fit indices, being an integral part of the SEM process, have a short but rabid history. Initially, Chi-Square tests were used; however, the test was proved ineffectual due to the large sample sizes that are required for SEM analysis (Gullen, 2000). As a result, various indices were developed to supplement the model fit analysis (Bollen, 1989). Fit indices can be classified into two categories – Model Test Statistic and Approximate Fit Index (Kline, 2011).

Model Test Statistics and Chi-Squared

In the model test statistic, the model data are compared to a baseline model. In this case, the baseline model is a covariance matrix of a sample of the data. If the covariance matrix of the overall data matches the covariance matrix of the sample, the model is considered a good fit. If the matrixes differ, the discrepancies using the model need to be explained. (Kline, 2011)

Model test statistics are typically developed as a "badness-of-fit" (Kline, 2011, p. 193) test. This means that failure to reject the null hypothesis indicates a good fit. Therefore, it is preferable for the resultant model test statistic to be as small as possible. Kline (2011) indicated that badness-of-fit tests are weak, as "lack of evidence to disprove an assertion…does not prove that the assertion is true." (p. 294).

The most basic model test statistic is the Model Chi-Square test. This test was developed by Karl Pearson (1900) and has withstood the test of time. It is probably the most well-known and accepted fit statistic. Its value lies in that it is nonparametric. The formula, according to Neave and Worthington (1988) is:

$$\chi^2 = \Sigma \frac{(Observed - Expected)^2}{Expected}) \quad \text{Eq. (3)}$$

Therefore, the Chi-Squared statistic is a percentage of the squared deviation from the expected over the expected score. A large Chi-Squared statistic indicates a large deviation from the expected distribution. The critical score is based on a Chi-Squared distribution table, Table 1 in the Appendix A.

In SEM computer software programs, the Chi-Squared test statistic is used to test for multivariate normality. However, as stated above, the basis of the test itself is nonparametric. This test can therefore be used to determine how well the SEM fit indices reflect the fit of the samples to the expected distribution, even when the data are expected to be non-normal.

Although the Chi-Squared test is nonparametric, the test is subject to other factors that can affect the reliability of the Chi-Squared test results. Reliability in SEM refers to the "ability of an item to perform a required function under stated conditions for a stated period of time" (Naresky, 1970, p. 199) and the probability that a test will perform or provide results according to its specified function (Naresky, 1970). In essence, it is the numerical value of the correlation coefficient. In other words, a test is considered reliable if it provides consistent results with consistent input values. A test that is not reliable would result in various values, even when situations do not differ.

A common measure of reliability is the Pearson Correlation Coefficient. The formula according to Hinkle, Wiersma, and Jurs (1998) is:

$$r_{xy} = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{[n\Sigma X^2 - (\Sigma X)^2] \times [n\Sigma Y^2 - (\Sigma Y)^2]}} \qquad \text{Eq. (4)}$$

In the equation above, the sample size value 'n' is included in the denominator. Therefore, reliability is sensitive to attenuation. Also, larger variances are better 'tolerated' when calculated reliability for data sets of large sample sizes. Depending on the application, a test may considered reliable when the calculated reliability coefficient is at least 0.7 or higher (indicating a strong correlation).

Factors that can affect the reliability of the SEM Chi-Squared tests include large correlations among variables, unique variance, and large sample size (Kline, 2011). When observed variables are highly correlated, the Chi-Squared value tends to increase. Unique variances among variables, which are a result of score unreliability, result in a loss of statistical power. As the Chi-Squared test is a badness-of-fit test, the loss of power reduces the probability of determining a poor model fit. As indicated above, the Chi-Squared value tends to increase with sample size.

Approximate Fit Indices

The second type of fit statistic is the approximate fit index. The difference between approximate fit indices and model test statistics is that fit statistics are based on continuous measures. In other words, there is not a dichotomous conclusion to either reject or accept a null hypothesis. The value of the fit statistic, as it compares to an ideal value in magnitude, provides a representation of the fit. For example, the ideal value for CFI fit index is 1.0. A model resulting in a CFI of 0.90 would be a better fit than a model resulting in a CFI value of 0.85. As the null hypothesis is not rejected at a decided alpha value, the magnitude of the value has meaning. Therefore, these fit indices can be considered as “rules-of-thumb” (Kline, 2011, p. 197) as opposed to “golden rules” (Kline, 2011, p. 197).

Approximate fit indices do not “distinguish between what may be sampling error and what may be real covariance evidence against the model” (Kline, 2011, p. 195). Thus, they do not provide information in regards to specification error. These tests are typically goodness-of-fit tests, which mean the ideal index statistic occurs at a value of a specified magnitude (such as 1.0 as opposed to zero).

Approximate fit indices provide approximations of model fit in four ways – absolute, incremental or comparative, parsimony-adjusted, and predictive (Kline, 2011). Absolute fit indices approximate fit by determining what proportion of a sample covariance matrix is explained by the model, in a similar fashion as the coefficient of determination (R^2). Incremental fit indices determine fit by comparing the model fit to a baseline model of zero population covariances among the variables. As Miles & Shevlin (2007) succinctly signified, “The indices effectively say ‘How well is my model doing, compared with the worst model that there is?’” (p. 870). Parsimony-adjusted indices provide a penalty for complex models. Predictive fit indices compare model fit to a sample group of randomly selected replicated samples based on the population. These categories are not mutually exclusive. An approximate fit index can fall into one or multiple of the categories indicated above. The most common of the approximate fit indices are RMSEA, SRMR and CFI. These tests are described in further detail below.

Root Mean Square Error Approximation (RMSEA)

The RMSEA is a parsimony-adjusted index. It is not a measure of central tendency but follows a non-central Chi-Square distribution. As such, the RMSEA has a high and a low value that are provided by most SEM software. The RMSEA is a badness-of-fit test, therefore a good fit indicator occurs when the RMSEA low value is less than 0.05 and the high value is less than 0.10. (Kline, 2011).

As a parsimony-adjusted index, the RMSEA adjusts for parsimonious characteristics. The mathematical construct of the index is divided by degrees of freedom of the SEM model. Kline (2011) specified the mathematical construct of the index.

$$RMSEA = \sqrt{\frac{\chi^2_M - df_M}{df_M(N-1)}}, \qquad \text{Eq. (5)}$$

where df_M = degrees of freedom of the SEM, N = sample size, and $\chi^2{}_M$ = Chi-Squared statistic value.

A small Chi-Squared value would indicate a good model fit. A model with a large degree of freedom, or a parsimonious model, would result in a small RMSEA value. In other words, parsimonious models that have small deviations would indicate a good model fit per this index. The equation is further divided by the sample size. Therefore, the parsimonious effect of the equation increases as sample size increase.

The limitations of RMSEA are obvious. The index contains inherent prejudices towards models that have large sample sizes and large degrees of freedom. A model with a moderate to large variation from the expected values, but with a large sample size, could pass the RMSEA criteria for model fit.

Standardized Root Mean Square Residual (SRMR)

Although the name is similar to the RMSEA, the two indices are quite different (Iacobucci, 2009). As per its name, the SRMR is a measure of the standardized value of the square root of the mean absolute covariance squared residual. As such, a good fit value would be close to zero. Hu and Bentler (1999) indicated that a maximum allowable value for a good fit would be approximately 0.09. Kline (2011) argued that this number is large, and can allow for large deviations. “This is because if the average absolute correlation residual is around 0.08, then many individual values could exceed this value, which would indicate poor explanatory power at the level of pairs of observed variables” (p. 209). Therefore, Kline (2011) recommended reviewing the matrix of correlation residuals and describing the patterns in a report, as opposed to merely stating the fit statistic.

The mathematical construct for the SRMR is provided by Iacobucci (2009) and Schermelleh-Engel, Moosbrugger, & Muller (2003).

$$SRMR = \sqrt{\frac{\Sigma_{i=1}^{p}\Sigma_{j=1}^{i}\left[\frac{(s_{ij}-\hat{\sigma}_{ij})}{(s_{ii}s_{jj})}\right]^2}{\frac{k(k+1)}{2}}}, \text{ Eq. (6)}$$

where k = observed endogenous variables + observed exogenous variables. S_{ij}, S_{ii}, and S_{jj} = value from the covariance matrix, and σ_{ij} = value from the expected matrix covariance.

Comparative Fit Index (CFI)

The CFI is an incremental fit index and a parsimony-adjusted index, where the data set is compared to the Chi-Squared values of a baseline model. This test performs well, even with small sample sizes. This is a goodness-of-fit test where a value of ‘one’ indicates the best fit. CFI was developed with the assumption that latent variables are not correlated (Hooper, Coughlan, and Mullen, 2008). Therefore, models with highly correlated latent variables can result in an inaccurate assessment of model fit.

The mathematical construct is a function of the Chi-Squared value and degrees of freedom of the model. The equation, as provided by Kline (2011) is written below:

$$CFI = 1 - \frac{\chi^2_M - df_M}{\chi^2_B - df_B}, \quad \text{Eq. (7)}$$

where df_X = degrees of freedom of the SEM/Baseline models, ${\chi^2}_X$ = Chi-Squared statistic value for the SEM/Baseline models, M = SEM model, and B = Baseline model.

This equation results in higher values for models with larger degrees of freedom, resulting in a more favorable fit statistic. Hu & Bentler (1999) provided a minimum recommended CFI value of 0.95 for acceptable fit criteria.

Fit Indices in Computer Software

Although the model fit indices provided by SEM computer software establish normality as the baseline condition (Arbuckle, 2011), the actual formulas for calculating the model fit numerical value are apparently nonparametric. It would therefore be reasonable to assume that using appropriate baseline or expected model distributions, the model fit index equations could be used to assess model fit for any distribution. However, the robustness of the formulas have not yet been assessed, and the capability of the indices to measure model fit for physical science data is of great interest.

CHAPTER 3 METHODOLOGY

Procedures

The purpose of this study is to evaluate the ability of the SEM methodology to obtain an accurate model for the physical sciences. SEM methodology was developed for the social and behavioral sciences. Typical software, such as Amos Graphics, is programmed with normality as a baseline condition, although fit indices apparently do not require normality. It is the intent of this study to determine whether common fit indices (e.g. RMSEA, SRMR, & CFI) provide accurate estimations of fit when data are sampled from non-normal distributions, as is typical for the physical sciences.

The validity of model fit indices will be analyzed for three study conditions. They are (1) mathematical distributions, (2) theoretical physical science distributions, and (3) applied physical science data sets.

(1) The mathematical distributions will be based on the logarithmic and sine wave distributions. These are commonly encountered in the physical science.

(2) The theoretical physical science distribution will be based on an equation developed for modeling fragment size distribution and was developed by Cheong, Reynolds, Salman, & Hounslow (2004). The equation of distribution is written below:

$$y = 1 - exp\left[-\left(\frac{x}{x_c}\right)^m\right], \qquad \text{Eq. (8)}$$

Where y = volume-fraction of fragment size of x, m, and x_c; m = distribution width; and x_c = fragment size ratio of 1-1/e.

The mathematical equation above will be limited to the values indicated in Figure 13.

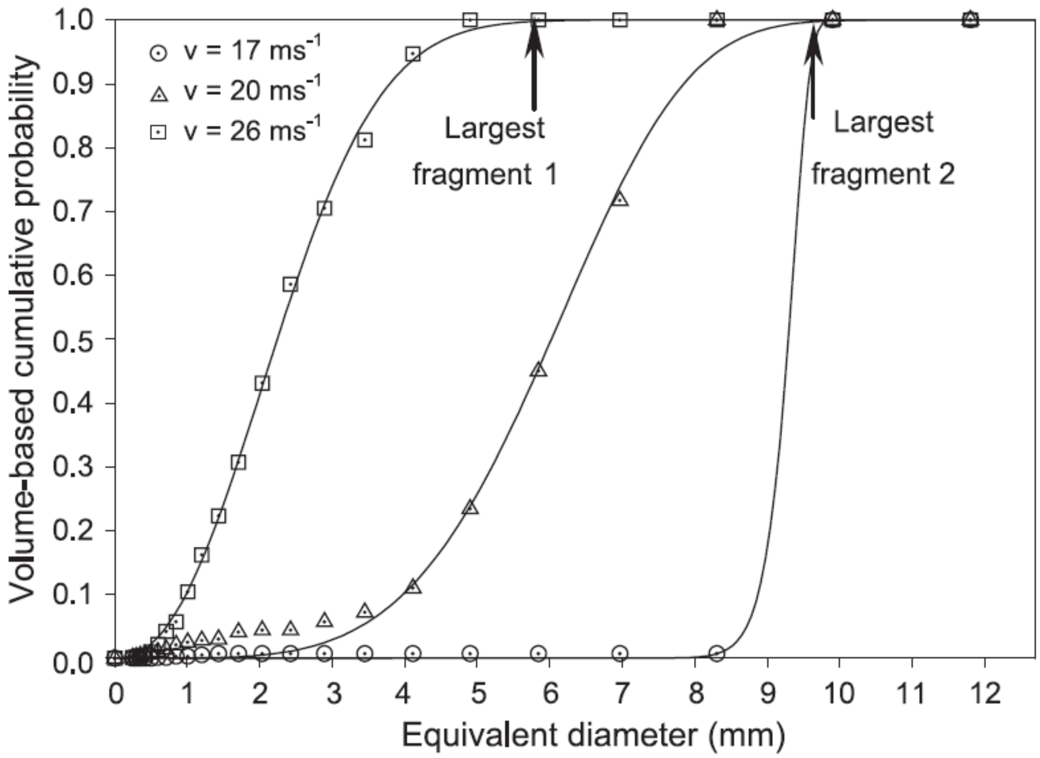

Figure 13. Glass Fragments Cumulative Size Distributions Based on Impact Velocity [Taken from Cheong, Reynolds, Salman, & Hounslow (2004, p. 230)].

(3) The applied physical science data sets will consist of real data, obtained through physical science research. The Shapiro Wilk and/or Kolmogorov-Smirnov tests will be used to determine if the distributions of the data sets are normal.

Data sets for study conditions 2 and 3 listed above will be obtained or developed using the mathematical formula indicated above. Samples from each of the data sets for all three study conditions will be obtained using a set seed value and pseudo-random number generator. This process will be repeated 10,000 times and the results analyzed using a Monte Carlo simulation. Each sample set will be analyzed using the Chi-Squared, and then using the RMSEA, SRMR, and CFI fit indices.

As the sample sets for study types 1 and 2 will be obtained from a mathematical formula of known distributions, the fit indices for the randomly selected samples should indicate a perfect model fit. Therefore, the model fit indices will be assessed to determine how closely the resultant values represent a perfect model fit.

As the sample sets for study type 3 will be obtained from real data, the sample sets will not have an exact known distribution. Therefore, assessment of model fit performance will be based on the correlation of the model fit indices in assessing a poor or good model fit. To this purpose, the model fit index test results will be standardized such that a value of 1.0 will equal a perfect model fit for all fit index tests. If the test results have a high correlation (value of 0.7 or greater), the model fit indices will be considered a good measure of model fit for non-normal distributions. If the results have a low correlation (value of 0.3 or lower), the model fit indices would be considered a poor measure of fit for non-normal distributions.

Identification of Extraneous Variables/Sources of Errors

1. Only 3 out of possibly dozens of fit indices will be used in this study. It is possible that other fit indices would provide different results.
2. The Chi-Squared Test is the standard for identifying model fit. If the test results in incorrect estimations (as typically happens in one out of twenty cases when the alpha value is set to 0.05), then the estimation of the validity of the model fit indices could be incorrect.

Sampling Plan

A pseudo-random number generator with a set seed value will be used to select the samples from the theoretical distributions and applied data sets. Fortran and RStudio both contain commands for random sampling of from a data set.

Data Gathering Methods

Data will be gathered from existing physical science research projects. No new data will be collected.

Data Analysis Software

Data analysis software will include Fortran 90 programming software with Compaq Visual Fortran 6.6c, RSTudio, and Notepad++ software programs.

Input Data Format

The input data will be formatted as metric (ratio) numbers of finite values. Nominal and Ordinal data are not typically used in the physical sciences.

Statistical Tests

Statistical tests and fit indices used to analyze the model fit for the SEM model are the Chi-Squared Test and the RMSEA, SRMR, and CFI fit indices.

Statistical Hypotheses

The null hypothesis for these tests will depend on whether the prospective test is a goodness-of-fit or badness-of-fit test. The null hypothesis for a goodness-of-fit test is that the data are a poor fit. The null hypothesis of a badness-of-fit test is that the data are a good fit.

Nominal Alpha

The alpha level for all statistical tests will be a value of 0.05. This value is standard for many physical science studies (Guralnik et al., 2000; Nichols, Morgan, Chabot, Sallis, & Calfas, 2000; and Petit et al., 2004).

Description of Computation Method

The applied data (study type 3) will be uploaded into Fortran 90 or RSTudio with the Lavaan library package to assess the frequency distribution and variation from normality. The data from the theoretical physical science distribution (study type 2) will be input into Fortran 90 or RStudio to develop the data set based on the mathematical formulas with a sample size sufficient for samples to be taken. Samples for all three study types will be randomly selected using Fortran 90 or RStudio with IMSL subroutines (indicated above). Estimation of model fit indices and Chi-Squared test values will be computed using Fortran 90 or RStudio. Computation of the correlation between the Chi-Squared test values and the model fit indices will be conducted using either Fortran 90 programming software with Compaq Visual Fortran 6.6c or RStudio.

Presentation of Results

The results will be presented in tabular and graphical format. A graphic showing the real and hypothetical data frequency distribution will be provided. Tables indicating Chi-Square, fit indices, and correlation values will also be provided. An analysis of the results, including assessment of the correlation values described above, will be provided in a written format.

CHAPTER 4 RESULTS

Due to unintended findings, discussed below, the presentation of the results has changed from the order of presentation in Chapter 3. A data set containing five variables (Torque at Transmission, Engine Speed, Vehicle Speed, Accelerator Pedal Position, and Fuel Used) was provided by Ford Motor Company. These data were obtained as part of the third data set type, the applied physical science data set consisting of real data obtained through physical science research. The data set originally contained 908,900 data points. After cleaning the data by deleting all accelerator pedal position values of 0.00, the revised data set contained 181,524 data points used in the analyses.

Test of Normality

The variables were tested for normality using RStudio to access Revolution-R (R). The results were compiled in Table 4. Due to a limitation of RStudio, the Shapiro-Wilk test could only be performed on a data set with a maximum sample size of 5,000. Therefore, a Monte Carlo simulation of 50,000 iterations of randomly selected samples with a sample size of 100 was performed.

Table 4

Shapiro-Wilk Test of Normality

Torque at Transmission	Engine Speed	Vehicle Speed	Accelerator Pedal Position	Fuel Used
0.1620185	0.008070679	0.001354129	0.1450409	0.0132040

The distributions for three of the five variables were non-normal, with nominal alpha set to 0.05. Only Torque at Transmission and Accelerator Pedal Position were not statistically significantly different from a normal distribution. Hence, typical physical science data sets could have both normal and non-normal variables within the same analysis.

Monte Carlo Results for Model Fit – Four Variables

Six Monte Carlo simulations of varying sample sizes and with 10,000 iterations of a Structural Equation Model (SEM) with four variables were performed in RStudio using the Lavaan library package. Sample sizes for the Monte Carlo simulations were n = 10, 20, 30, 50, and 250. The exogenous variables in the SEM were Accelerator Pedal Position, Engine Speed, and Vehicle Speed. The endogenous variable was Torque at Transmission. The results of the analyses were summarized in Table 5. The full output from RStudio is presented in Figures B1 through B6 in Appendix B.

Model fit improved as sample size decreased, as indicated in Table 5. Indication of a poor model fit approximately 100% of the time for sample size of 250 and 88% of the time for sample size of 20 resulted when examining the trends for model fit index RMSEA Lower. Indication of a poor model fit approximately 91% of the time for sample size of 250 and approximately 84% of the time for sample size of 20 resulted when examining model fit index results for CFI. Indication of a poor model fit approximately 100% of the time resulted when examining the results for the other model fit indices (Chi-Squared, RMSEA Upper, and SRMR).

Table 5

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Sample Size	250	100	50	30	20	10
Mean Correlation	0.46	0.43	0.41	0.40	0.39	0.04
Mean Degrees of Freedom	3	3	3	3	3	0.34*
Chi-Squared	100%	100%	100%	100%	100%	N/A**
RMSEA Lower	100%	99%	95%	91%	88%	N/A**
RMSEA Upper	100%	100%	100%	100%	100%	N/A**
SRMR	100%	100%	99%	99%	100%	N/A**
CFI	91%	83%	81%	83%	84%	N/A**

*Result for mean degrees of freedom for 10,000 iterations was not a whole number. This implies that the degrees of freedom varied between repetitions.
**The results for model fit did not reflect a logical reality. (i.e. The percentage of time that RMSEA Lower was less than 0.05 was eleven percent and the percentage of time that RMSEA Lower was greater than 0.05 was ten percent. These values do not add to 100%)

Correlation decreased as sample size decreased; from r = 0.46 for a sample size of n = 250 to r = 0.39 for a sample size of n = 20. As the correlation approached zero (e.g., sample size of n = 10), the results for the model fit from the Monte Carlo simulation resulted in illogical values. The percentages of greater than and less than critical values did not result in percentage numbers that added to 100%. This is shown in greater detail in Figure 14.

sample size=10
repetitions=10000
mean correlation=0.0421997270262573
mean DOF*=0.3432
chisq a**<0.05=0.1144-> This means that p-chi-squared was less than 0.05 11% of the time
chisq a**<0.01=0.1143-> This means that p-chi-squared was less than 0.01 11% of the time
RMSEA (lower)**<=0.05=0.0186-> This means that RMSEA Lower is less than 0.05 2% of the time
RMSEA (lower)**>0.05=0.0958-> This means that RMSEA Lower is greater than 0.05 10% of the time
RMSEA (upper)**<=0.1=0-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)**>0.1=0.1144-> This means that RMSEA Upper is greater than 0.1 11% of the time
SRMR**<=0.09=0-> This means that SRMR is less than 0.09 0% of the time
SRMR**>0.09=0.1144-> This means that SRMR is greater than 0.09 11% of the time
CFI**<0.9=0.0994-> This means that CFI is less than 0.9 10% of the time
CFI**<0.75=0.0766-> This means that CFI is less than 0.75 8% of the time
CFI**<0.5=0.0413-> This means that CFI is less than 0.5 4% of the time

*The mean degrees of freedom did not result in a whole number, implying varying degrees of freedom between repetitions.
**The results for model fit did not reflect a logical reality. (i.e. The percentage of time that RMSEA Lower was less than 0.05 was eleven percent and the percentage of time that RMSEA Lower was greater than 0.05 was ten percent. These values do not add to 100%)

Figure 14. Lavaan Output for Sample Size of 10 and Four Variables, Repetitions = 10,000.

Monte Carlo Results for Model Fit – Five Variables

A Monte Carlo simulation was performed for a second data set with five variables. The second data set included the cleaned data from the four variable SEM analyses and a fifth variable of Fuel Used. The exogenous variables in the SEM were Accelerator Pedal Position, Engine Speed, Vehicle Speed, and Torque at Transmission. The endogenous variable was Fuel Used.

Initially, the intent was to perform six Monte Carlo simulations with sample sizes and iterations, the same as for the four variable SEM analyses. However, the results of the analysis with sample size of 250 were illogical, similar to those of the four variable analysis with sample size of n = 10 (refer to Figures 14 and 15). Further analyses demonstrated Fuel Used had a small correlation with the other four variables in the analysis, as noted in the correlation coefficient matrix compiled in Table 6.

Table 6

Correlation Matrix

Variable	Torque at Transmission	Engine Speed	Vehicle Speed	Accelerator Pedal Position	Fuel Used
Torque at Transmission	1.000	0.404	0.275	0.968	0.007
Engine Speed	0.404	1.000	0.727	0.391	0.022
Vehicle Speed	0.275	0.727	1.000	0.210	0.030
Accelerator Pedal Position	0.968	0.391	0.210	1.000	0.007
Fuel Used	0.007	0.022	0.030	0.007	1.000

Sample Size=450
repetitions=10000
mean correlation x1&z=0.000462070478175039
mean correlation x2&z=0.0025020921013811
mean correlation x3&z=0.00227332378855168
mean correlation x4&z=0.000598797833551206
mean DOF= numeric(0)*
chisq a**<0.05=0.0325-> This means that p-chi-squared was less than 0.05 3% of the time
chisq a**<0.01=0.0325-> This means that p-chi-squared was less than 0.01 3% of the time
RMSEA (lower)**<=0.05=0-> This means that RMSEA Lower is less than 0.05 0% of the time
RMSEA (lower)**>0.05=0.0325-> This means that RMSEA Lower is greater than 0.05 3% of the time
RMSEA (upper)**<=0.1=0-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)**>0.1=0.0325-> This means that RMSEA Upper is greater than 0.1 3% of the time
SRMR**<=0.09=0-> This means that SRMR is less than 0.09 0% of the time
SRMR**>0.09=0.0325-> This means that SRMR is greater than 0.09 3% of the time
CFI**<0.9=0.0325-> This means that CFI is less than 0.9 3% of the time
CFI**<0.75=0.0325-> This means that CFI is less than 0.75 3% of the time
CFI**<0.5=0.0325-> This means that CFI is less than 0.5 3% of the time

*RStudio was not able to calculate the degrees of freedom.

** Results indicated that RMSEA Lower was less than 0.05 zero percent of the time and greater than 0.05 three percent of the time. These values do not add up to 100%.

Figure 15. Lavaan Output for Sample Size of 250 and Five Variables, Repetitions = 10,000.

Second Analysis – Determine Minimum Correlation Coefficient for SEM

A mutual characteristic exhibited by the five variable analysis and the four variable analysis with sample size of n = 10 is the low correlation values associated with the model. Apparently, when correlation was low the results of the model fit were inexplicable. This unintended outcome led to conducting additional research to determine what constituted a low correlation, and whether there was a minimum allowable correlation value between variables that is a prerequisite for a SEM to be meaningful.

Hence, additional Monte Carlo simulations were conducted. Based on 1,000 repetitions, further analyses included an SEM model of four variables with input values consisting of a correlation matrix of numbers of randomly selected values limited to a specified range. An indication of the legitimacy of the model fit indices would be an indication of poor model fit, because the input correlation values for the four variables had no relationship and were randomly selected. An indication of good model fit would signify an illegitimate measurement.

The first Monte Carlo simulation included a correlation matrix of random values from a range of 0.04 plus or minus 0.015. All model fit indices results included in the analyses (Chi-Squared, RMSEA Lower, RMSEA Upper, SRMR, and CFI) were an indication of a poor model fit 0% of the time. Refer to Table 7 below.

Table 7

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.04 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	0%	0%
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	0%	0%	0%	0%	0%	0%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	0%	0%	0%	0%

As the correlation matrix values were increased in magnitude, the results of the model fit indices became illogical. The output values for the model degrees of freedom varied between repetitions. Furthermore, the percentages of greater than and less than critical values did not result in percentage numbers that added to 100%. These illogical results are similar to those that occurred in the four and five variable SEM analyses indicated in the sections above. The model fit indices results ceased to be illogical as the correlation magnitudes were continuously increased, and instead the results were an indication of a poor model fit with increasing reliability. At a certain correlation magnitude (e.g. when correlation was equal to 0.08 plus or minus 0.015 as in Table 11), the results of the model fit indices were an indication of a poor model fit for the conditions studied for all Monte Carlo repetitions. The same applies to Tables 11, 19, 21, 27, and 30.

A summary of these results from the Monte Carlo simulations for the varying correlation magnitudes and sample sizes are provided in Tables 8-30. The increasing reliability of the different model fit index tests was clearly indicated.

Table 8

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.05 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	0%	0%
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	0%	N/A	N/A	N/A	N/A	N/A
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	0%	0%	N/A	N/A

Table 9

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.06 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	0%	N/A
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	35%	N/A	N/A	N/A	N/A	N/A
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	0%	N/A	N/A	N/A

Table 10

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.07 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	0%	N/A
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	92%	N/A	N/A	N/A	N/A	N/A
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	0%	N/A	N/A	N/A

Table 11

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.08 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	N/A	N/A	N/A	N/A

Table 12

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.09 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	N/A	N/A	N/A	N/A

Table 13

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.10 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	0%	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	0%	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	N/A	N/A	N/A	N/A	N/A

Table 14

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.11 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	0%	N/A	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	0%	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	N/A	N/A	N/A	N/A	N/A

Table 15

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.12 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	N/A	N/A	N/A	N/A	N/A

Table 16

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.13 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	N/A	N/A	N/A	N/A	N/A

Table 17

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.14 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	0%	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	0%	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	9%	N/A	N/A	N/A	N/A	N/A

Table 18

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.15 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	0%	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	2%	N/A	N/A	N/A	N/A	N/A
CFI	75%	N/A	N/A	N/A	N/A	N/A

Table 19

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.16 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	54%	N/A	N/A	N/A	N/A	N/A
CFI	100%	100%	100%	100%	100%	100%

Table 20

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.17 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	N/A	N/A	NA	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	99%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 21

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.18 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 22

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.19 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 23

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.20 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 24

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.21 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 25

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.22 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	4%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 26

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.23 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	86%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 27

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.24 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	100%	100%	100%	100%	100%	100%
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 28

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.25 Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Chi-Squared	100%	100%	100%	100%	100%	100%
RMSEA Lower	24%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 29

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.26 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	100%	100%	100%	100%	100%	100%
RMSEA Lower	11%	N/A	N/A	N/A	N/A	N/A
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 30

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range of 0.27 Plus or Minus 0.015

Sample Size	50	100	150	200	300	500
Chi-Squared	100%	100%	100%	100%	100%	100%
RMSEA Lower	100%	100%	100%	100%	100%	100%
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Each model fit index resulted in legitimate results at different correlation magnitudes, as indicated in Tables 7 through 30 above. The best model fit index, which resulted in legitimate model fit estimation at the lowest correlation magnitude, was RMSEA Upper at a correlation of 0.08 for all sample sizes. The next best model fit index was CFI, with valid estimation of model fit at a minimum correlation value of 0.16. The next best model fit index following CFI was SRMR, with valid model fit estimation at a minimum correlation value of 0.17 for large sample sizes and 0.18 for sample size of 50. The next best model fit index following SRMR was Chi-Squared, with valid model fit estimation at a minimum correlation value of 0.24. The model fit index that performed the poorest was RMSEA Lower, with valid model fit estimation at a minimum correlation of 0.27. These results are presented in greater detail in Table 31 below.

Table 31

Minimum Correlation Values for Valid Model Fit Index Measurement

	Sample Size					
Model Fit Index	50	100	150	200	300	500
Chi-Squared	0.24	0.24	0.24	0.24	0.24	0.24
RMSEA Lower	0.27	0.27	0.27	0.27	0.27	0.27
RMSEA Upper	0.08	0.08	0.08	0.08	0.08	0.08
SRMR	0.18	0.17	0.17	0.17	0.17	0.17
CFI	0.16	0.16	0.16	0.16	0.16	0.16

CHAPTER 5 CONCLUSIONS

The purpose of this study was to evaluate the sensitivity of selected fit index statistics in determining model fit when the distribution varied from normality, as is typically true of data research for the physical sciences. SEM is a popular statistical method and is used in many physical and social behavioral science research projects; however, the sensitivity of the model fit indices when normality is violated had never been estimated. Estimation of the model fit is the first and an essential part of estimating the SEM, it was therefore imperative to study the effects when normality was violated.

The equations for calculating the model fit indices were apparently nonparametric. It was hypothesized that using appropriate baseline or expected model distributions, the model fit index equations could be used to assess model fit for any distribution. It was expected that these equations would be robust in calculating model fit.

The original intent for performing the research was to analyze the legitimacy of the model fit indices' results against three different types of distributions. These distribution types were (1) mathematical distributions, (2) theoretical physical science distributions, and (3) applied physical science data sets. These data sets were sufficiently large to serve as proxies for typical physical science distributions and were meant to be analyzed using the specified model fit index tests (Chi-Squared, RMSEA Upper, RMSEA Lower, SRMR, and CFI).

Due to unintended results, the research findings from the third data set (applied physical science data set) were presented first. This data set contained five variables (Torque at Transmission, Engine Speed, Vehicle Speed, Accelerator Pedal Position, and Fuel Used). The variables were tested for normality using RStudio. The distributions for three of the five variables were non-normal, with nominal alpha set to 0.05. Only Torque at Transmission and Accelerator Pedal Position were not statistically significantly different from a normal distribution. Hence, typical physical science data sets could have both normal and non-normal variables within the same analysis.

The variables were assessed using an SEM, and a Monte Carlo simulation of 10,000 iterations for varying sample sizes of n = 10, 20, 30, 50, and 250. It was determined that an indication of poor model fit occurred with greater consistency as the sample size of the data set increased. Indication of a poor model fit approximately 100% of the time for sample size of 250 and 88% of the time for sample size of 20 resulted when examining the trends for model fit index RMSEA Lower. Indication of a poor model fit approximately 91% of the time for sample size of 250 and approximately 84% of the time for sample size of 20 resulted when examining model fit index results for CFI. Indication of a poor model fit approximately 100% of the time resulted when examining the results for the other model fit indices (i.e., Chi-Squared, RMSEA Upper, SRMR).

It was also determined that the magnitude of the correlation decreased as sample size decreased; from r = 0.46 for a sample size of n = 250 to r = 0.39 for a sample size of n = 20. As the correlation approached zero (e.g., sample size of n = 10), the model fit from the Monte Carlo simulation resulted in illogical values.

The first indication of illegitimate results was the fractional value of 0.34 for the mean degree of freedom for the 10,000 repetition Monte Carlo simulation (Table 32). This value should be a whole number, and should be consistent for the same SEM. Although not the focus of this study, the fractional values indicated varying degree of freedom with each repetition of the same experiment.

The second indication of illegitimate results was the value indicating the percentage of time that the model fit indices were an indication of a poor model fit. The percentage of time the RMSEA Lower tests were greater than 0.05 was 10% and the percentage of time the RMSEA Lower tests were less than or equal to 0.05 was 2% (Table 32). These two values should add to 100%, not 12%. This illogical result only occurred for the Monte Carlo simulation for sample size n = 10. The larger sample size Monte Carlo simulations resulted in values that added to 100% (e.g., when sample size equaled to 250 as indicated in Table 32; the same applies to remaining Figures B2 – B6 in Appendix B).

The illogical results occurred again when a fifth variable was added to the SEM. Lavaan could not compute the degrees of freedom for this simulation, and the percentage of time model fit indices were less than or greater than a critical value did not sum to 100% (Table 32).

Table 32

Lavaan Output for Several Monte Carlo Simulations, Repetitions = 10,000

Variables	4	4	5
Sample Size	250	10	250
Degree of Freedom	3	0.3432	N/A
% Time RMSEA Lower <= 0.05	0%	19%	0%
% Time RMSEA Lower > 0.05	100%	10%	3%
% Time RMSEA Upper <= 0.1	0%	0%	0%
% Time RMSEA Upper > 0.1	100%	11%	3%
% Time SRMR <= 0.09	0%	0%	0%
% Time SRMR > 0.09	100%	11%	3%

These unexpected illogical results led to checking the Monte Carlo simulation command file for programming errors. The Monte Carlo command file was executed in RStudio, with one repetition of the SEM and a manufactured correlation matrix with six variables. The same correlation matrix was then used in IBM's SPSS Amos Graphics (version 22) with the same variables and the same SEM layout. The results from Amos Graphics were within rounding error of the Monte Carlo simulation in RStudio (Table 33 and Figures 16 and 17).

Table 33

Manufactured Correlation Matrix for Checking Lavaan Monte Carlo Programming File

type	varname	X1	X2	X3	X4	X5	X6
n		450	450	450	450	450	450
corr	X1	1					
corr	X2	0.084	1				
corr	X3	-0.072	-0.093	1			
corr	X4	0.050	-0.006	-0.010	1		
corr	X5	0.076	0.014	-0.039	-0.060	1	
corr	X6	-0.028	-0.069	-0.047	-0.097	0.023	1
stddev	1	1	1	1	1	1	1

```
lavaan (0.5-19) converged normally after  10 iterations
  Number of observations                       450
  Estimator                                    ML
  Minimum Function Test Statistic              7.008
  Degrees of freedom                           4
  P-value (Chi-square)                         0.135

User model versus baseline model:
  Comparative Fit Index (CFI)                  0.641
  RMSEA                                        0.041
  90 Percent Confidence Interval               0.000  0.090
  P-value RMSEA <= 0.05                        0.549
  SRMR                                         0.027

Regressions:
                        Estimate    Std.Err    Z-value    P(>|z|)
  X6 ~
    X1                  -0.028      0.047      -0.604     0.546
    X4                  -0.095      0.047      -2.019     0.044
    X3                  -0.049      0.047      -1.054     0.292
    X5                  0.018       0.047      0.382      0.703
  X5 ~
    X4                  -0.060      0.047      -1.275     0.202
    X3                  -0.039      0.047      -0.826     0.409
    X2                  0.010       0.047      0.208      0.835
  X3 ~
    X2                  -0.093      0.047      -1.982     0.047
  X4 ~
    X1                  0.050       0.047      1.059      0.289
    X3                  -0.006      0.047      -0.138     0.890
  X1 ~
    X2                  0.084       0.047      1.782      0.075

Variances:
               Estimate    Std.Err    Z-value    P(>|z|)
    X6         0.985       0.066      15.000     0.000
    X5         0.993       0.066      15.000     0.000
    X3         0.989       0.066      15.000     0.000
    X4         0.995       0.066      15.000     0.000
    X1         0.991       0.066      15.000     0.000
```

Figure 16. Lavaan Output for Sample Size of 450 and Six Variables.

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
X3	<---	X2	-.093	.047	-1.980	.048	par_2
X1	<---	X2	.084	.047	1.780	.075	par_3
X4	<---	X1	.050	.047	1.058	.290	par_4
X4	<---	X3	-.006	.047	-.138	.890	par_10
X5	<---	X2	.010	.047	.208	.835	par_1
X5	<---	X3	-.039	.047	-.825	.409	par_8
X5	<---	X4	-.060	.047	-1.273	.203	par_9
X6	<---	X1	-.028	.047	-.603	.546	par_5
X6	<---	X4	-.095	.047	-2.016	.044	par_6
X6	<---	X3	-.049	.047	-1.052	.293	par_7
X6	<---	X5	.018	.047	.381	.703	par_11

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
X2	.998	.067	14.983	***	par_12
d1	.989	.066	14.983	***	par_13
g1	.991	.066	14.983	***	par_14
a1	.995	.066	14.983	***	par_15
e1	.993	.066	14.983	***	par_16
c1	.985	.066	14.983	***	par_17

Model Fit Summary

Chi-Squared

Model	NPAR	Chi-Squared	DF	P	CMIN/DF
Default model	17	6.992	4	.136	1.748

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.700	-.124	.845	-.348	.640

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.041	.000	.090	.549
Independence model	.035	.000	.062	.801

Figure 17. Amos Output for Sample Size of 450 and Six Variables.

In an effort to simplify the SEM to better understand the source of the issues, a Monte Carlo simulation with 10,000 repetitions of varying correlation matrices of randomly selected numbers between -0.1 and +0.1 was executed in RStudio. The model fit results of the Monte Carlo simulation were still illogical, similar to the illogical results for the four variable analysis with sample size n = 10 and for the five variable analyses. The output of the latest repetition of the Monte Carlo simulation was extracted and compared with the output from Amos Graphics. They were the same within rounding error.

Fit index results should be consistent, regardless of whether a meaningful model is produced. Examination of the model fit indices results should indicate a good or a poor model fit when a reasonable model is assessed. However, examination of the results should never indicate a good model fit on a poorly defined model. In this case, the correlation values between variables were small and the paths were not significant. Therefore the model, having no relationships, should result in an indication of poor model fit when assessed using model fit index tests. This indication of poor model fit should occur uniformly for all model fit index tests and for all sample sizes.

These findings were discussed with statistical experts. The first subject matter expert believed there were some insights that could be garnered based on the results (B. Zumbo, personal communication, November, 2015). The second subject matter expert confirmed the Lavaan Monte Carlo command file by executing the model in Mplus (Figure 18) using the manufactured correlation matrix (Table 33), and noted:

> [T]his model shows mixed results in terms of fit, but the paths are not significant and the amount of explained variance is not significant – so you have a well-fitting model that explains zero variance. So in essence, the structural relations specified in the model were able to accurately reproduce a covariance matrix with no significant covariances. To say one needs to approach that with caution is a considerable understatement. (R. Partridge, personal communications, November, 2015).

As a beginning to approaching the model fit indices assessment with caution, additional research was conducted to determine what SEM conditions caused the model fit index results to be illogical. The Monte Carlo simulation models were assessed to find common characteristics. A consistent attribute between the five variable analysis and the four variable analysis for sample size n = 10 was the low correlation values between the variables. It appeared when the correlation values between variables were low the results of the model fit indices were illogical. Additional research was therefore conducted to determine what constituted a low correlation, and whether there was a minimum allowable correlation value between variables that is a prerequisite for a SEM to be meaningful.

Mplus VERSION 5.1

Number of continuous latent variables	0	Number of observations	450
Number of dependent variables	4	Number independent variables	2

TESTS OF MODEL FIT

Chi-Square Test of Model Fit

Value	7.749
Degrees of Freedom	5
P-Value	0.1706
SRMR Value	0.025
CFI	0.537

RMSEA (Root Mean Square Error Of Approximation)

Estimate	0.035
90 Percent C.I.	0.000 0.080
Probability RMSEA <= .05	0.649

MODEL RESULTS (Two-Tailed)

	Estimate	S.E.	Est./S.E.	P-Value
X4 ON X1	0.054	0.047	1.145	0.252
X3 ON X2	-0.093	0.047	-1.986	0.047
X5 ON				
X2	.010	0.047	0.216	0.829
X3	-0.037	0.047	-0.778	0.437
X6 ON				
X1	-0.030	0.047	-0.631	0.528
X1 WITH X2	0.083	0.047	1.754	0.079
X6 WITH				
X4	-0.095	0.047	-2.015	0.044
X5	0.024	0.047	0.513	0.608
X5 WITH X4	-0.063	0.047	-1.341	0.180
X3 WITH X4	-0.012	0.047	-0.265	0.791
Variances				
X1	0.998	0.067	15.000	0.000
X2	0.998	0.067	15.000	0.000
X3	0.989	0.066	15.000	0.000
X4	0.995	0.066	15.000	0.000
X5	0.996	0.066	15.000	0.000
X6	0.997	0.066	15.000	0.000

Figure 18. Mplus Output for Sample Size of 450 and Six Variables.

Additional Monte Carlo simulations were conducted, with 1,000 repetitions and varying magnitudes of correlation matrices. The magnitudes of the correlation values for the correlation matrix were randomly selected from a base value plus or minus 0.015. Twenty four Monte Carlo simulations were performed, with the base value increasing from 0.04 to 0.27 and a Monte Carlo simulation performed at every hundredths place value (i.e. 0.04, 0.05 ,0.06, etc.). An indication of the legitimacy of the model fit indices would be an indication of poor model fit, because the input correlation values for the four variables had no relationship and were randomly selected.

As the correlation matrix values were increased in magnitude, the results of the model fit indices became first illogical and then finally logical with an increasing indication of a poor model fit. At a certain correlation magnitude (e.g. when correlation was equal to 0.08 plus or minus 0.015 as in Table 34), the results of the model fit indices were an indication of a poor model fit for the model fit index studied for all Monte Carlo repetitions.

Each model fit index resulted in legitimate results at different correlation magnitudes. Model fit indices can be ranked from best to worst based on the minimum correlation values required before legitimate results were acquired. The model fit indices, from best to worst, are listed in Table 35 below with their respective minimum correlation values and ranks.

Table 34

Monte Carlo Simulation Percentage of Model Fit Indices Indication of Poor Model Fit

Correlation Matrix Magnitudes Range Plus or Minus 0.015						
Sample Size	50	100	150	200	300	500
Correlation	0.04	0.04	0.04	0.04	0.04	0.04
Chi-Squared	0%	0%	0%	0%	0%	0%
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	0%	0%	0%	0%	0%	0%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	0%	0%	0%	0%
Correlation	0.06	0.06	0.06	0.06	0.06	0.06
Chi-Squared	0%	0%	0%	0%	0%	N/A
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	35%	N/A	N/A	N/A	N/A	N/A
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	0%	N/A	N/A	N/A
Correlation	0.08	0.08	0.08	0.08	0.08	0.08
Chi-Squared	0%	0%	0%	0%	N/A	N/A
RMSEA Lower	0%	0%	0%	0%	0%	0%
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	0%	0%	0%	0%	0%	0%
CFI	0%	0%	N/A	N/A	N/A	N/A
Correlation	0.16	0.16	0.16	0.16	0.16	0.16
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	0%	N/A	N/A	N/A	N/A
SRMR	54%	N/A	N/A	N/A	N/A	N/A
CFI	100%	100%	100%	100%	100%	100%
Correlation	0.18	0.18	0.18	0.18	0.18	0.18
Chi-Squared	0%	N/A	N/A	N/A	N/A	N/A
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
SRMR	100%	100%	100%	100%	100%	100%
Correlation	0.24	0.24	0.24	0.24	0.24	0.24
Chi-Squared	100%	100%	100%	100%	100%	100%
RMSEA Lower	0%	N/A	N/A	N/A	N/A	N/A
Correlation	0.27	0.27	0.27	0.27	0.27	0.27
Chi-Squared	100%	100%	100%	100%	100%	100%
RMSEA Lower	100%	100%	100%	100%	100%	100%
RMSEA Upper	100%	100%	100%	100%	100%	100%
SRMR	100%	100%	100%	100%	100%	100%
CFI	100%	100%	100%	100%	100%	100%

Table 35

Minimum Correlation Values

Rank	Model Fit Index	Minimum Correlation Value
1	RMSEA Upper	0.08
2	CFI	0.16
3	SRMR	0.18
4	Chi-Squared	0.24
5	RMSEA Lower	0.27

The results from the last repetition of the Monte Carlo simulation with correlation range of 0.1 plus or minus 0.015 and sample size of 500 were extracted (refer to Table 36 and Figure 19 below) to better understand the results of the Monte Carlo simulations and to verify the conclusions determined above. The model fit was a mixture between good and poor, based on the results of the model fit index tests. The p value for the Chi-Squared test was 0.003, an indication of a poor model fit. The RMSEA Upper value was 0.133, an indication of a poor model fit. The RMSEA Lower value was 0.044, an indication of a good model fit. The CFI value was 0.505, an indication of a poor model fit. The SRMR value was 0.055, an indication of a good model fit.

Number of observations				500
Estimator				ML
Minimum Function Test Statistic				14.059
Degrees of freedom				3
P-value (Chi-square)				0.003
User model versus baseline model:				
Comparative Fit Index (CFI)				0.505
Tucker-Lewis Index (TLI)				0.010
Number of free parameters				7
RMSEA				0.086
rmsea.ci.lower				0.044
rmsea.ci.upper				0.133
90 Percent Confidence Interval				0.044 0.133
P-value RMSEA <= 0.05				0.075
SRMR				0.055
Parameter estimates:				
Information				Expected
Standard Errors				Standard
Regressions:				
	Estimate	Std.err	Z-value	P(>|z|)
z ~				
x1	0.088	0.044	1.990	0.047
x2	0.079	0.044	1.786	0.074
x3	0.098	0.044	2.232	0.026
Covariances:				
x1 ~~x2		0.000		
x3		0.000		
x2 ~~x3		0.000		
Variances:				
z		0.970	0.061	
x1		0.998	0.063	
x2		0.998	0.063	
x3		0.998	0.063	

x1=Accelerator Pedal Position, x2=Engine Speed, x3=Vehicle Speed, and x4=Torque At Transmission

Figure 19. Lavaan Output for Sample Size of 500 and Four Variables, Repetitions = 1,000.

Table 36

Correlation Matrix

Variables	z	X1	X2	X3
z	1.000	0.104	0.098	0.115
X1	0.104	1.000	0.100	0.088
X2	0.098	0.100	1.000	0.109
X3	0.115	0.088	0.109	1.000

The regression coefficients for the exogenous variables were 0.088 for X1, 0.079 for X2, and 0.098 for X3. Although these values were low, the coefficients for X1 and X3 were statistically significant. This is illogical, as the correlation magnitudes in the correlation matrix were all low. Statistically significant paths between variables are therefore a contradictory conclusion. These results solidified the conclusion above that a SEM with a correlation matrix of low values would result in illogical outcomes.

The results from Tables 34 and 35 were forwarded to a third subject matter expert who opined:

> SEM is a collection of procedures that are assessed based on a plethora of fit or lack of fit statistics that could be subjectively chosen or ignored to support or eliminate a model. Moreover, dozens of caveats (such as those listed in Kline, 2011, e.g., at its core it relates to non-experimental data and hence there can never be causation (p. 8), a poor model can be preserved by modifying the hypotheses on which it is based (p. 8), alternative models may not be ruled out (p. 8), it is a large sample technique (p. 11), it eschews hypothesis testing and hence is veiled behind subjectivity (p. 13), the statistical significance of estimated parameters are dependent on the algorithm adopted (p. 13), maximum likelihood estimate cannot tolerate even a single missing datum (p. 48), a nonpositive definite matrix cannot be

> analyzed (p. 49), ill-scaled covariance matrices cannot be handled (p. 67)) severely limit SEM outside of textbook examples. (S. Sawilowsky, personal communications, November, 2015).

This expert, in citing Kline (2011), also pointed to the most likely explanation for the study results:

> It may be problematic to submit for analysis just a correlation matrix without standard deviations or specify that all standard deviations are 1.0, which standardizes everything. This is because the default method of ML estimation (and most other methods, too) assumes that the variables are unstandardized. This means that if a correlation matrix without standard deviations is analyzed, the results may not be correct…Some SEM computer programs give warning message or terminate the run if the researcher requests the analysis of a correlation matrix only with standard ML estimation. By the same token, it would also be problematic to convert raw scores to z scores and then submit for analysis the data file of standardized scores. (Kline, 2011, p. 49)

The subject matter expert concluded that no systematic Monte Carlo study could be conducted by inputting an incrementally increasing correlation matrix, such as was attempted in this study (S. Sawilowsky, personal communications, November, 2015).

Recommendations for Future Research

Illogical values for percentage of times the model fit indices results were above or below certain values occurred when correlation values in the correlation matrix table were low. For example, the results from RStudio for percentage of time the RMSEA Lower tests were greater than 0.05 was 10% and the percentage of time the

RMSEA Lower tests were less than or equal to 0.05 was 2% (refer to Table 32 above). These two values should add to 100%, not to 12%.

The reason for these illogical results could be due to several factors. One reason could be due to a programming error from Lavaan and the error could be unique for that program. Alternatively, the illogical results could be due to an error in the SEM theory itself. At any rate, further research is required to determine the source.

An indication of illogical results, exhibited from the results of the Monte Carlo simulation for an SEM of four variables and sample size n = 10, was the fractional property of the mean degrees of freedom (for 10,000 repetitions) resulting in a value of 0.3432 (refer to Table 32 above). After this result occurred, the last of the 10,000 repetitions was assessed and the degree of freedom for that repetition was three, a whole number. Therefore, the fraction value for the mean degrees of freedom is due to a variant degree of freedom for each repetition. This is illogical. The degrees of freedom are a function of the variables, correlations, errors, and paths. If the number of correlations, errors, variables, or paths do not change then the degree of freedom should be uniform. The varying degree of freedom is illogical. Further research is required to determine the cause of this illogical result and the necessary boundary conditions to prevent disturbance of the SEM results.

Implications for Future Research

No systemic Monte Carlo study could be conducted by inputting an incrementally increasing correlation matrix, due to the many restrictions listed above. One possibility is to input a matrix of non-standardized values. This might be accomplished by incrementally increasing correlation values and using a covariance matrix of unstandardized values.

According to Nunnally and Bernstein (1994), the relationship between correlation and covariance is:

$$\sigma_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{N} = \frac{\sum xy}{N} \qquad \text{Eq. (9)}$$

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \qquad \text{Eq. (10)}$$

$$\sigma_{xy} = r_{xy}\sigma_x\sigma_y \qquad \text{Eq. (11)}$$

where N = number of observed variables, r_{xy} = correlation of x and y, σ_x = standard deviation of x, σ_y = standard deviation of y, and σ_{xy} = covariance of x and y.

A Monte Carlo simulation study might be designed by using covariance matrices developed from varying correlation and standard deviations values. If set to specific parameters, a minimum allowable standard deviation value for each correlation value can be determined. The parameters required to produce results are unknown, but an initial recommendation would be to run a Monte Carlo simulation of 1,000 repetitions; correlation values = 0.025, 0.05, 0.075, 0.1, etc.; sample size = 50, 100, 150, 200, 300, and 500; and standard deviations for x and y of different patterns (Table 37 below).

Table 37

*Varying Standard Deviation Values**

Pattern One		Pattern Two		Pattern Three		Pattern Four		Pattern Five		Pattern Six	
SD_x	SD_y	SD_x	SD_y	SD_x	SD_y	SD_x	SD_y	SD_x	SD_y	SD_x	SD_y
1	1.5	1	15	1.5	1.5	15	15	1.5	768	15	7680
1	3	1	30	3	3	30	30	3	384	30	3840
1	6	1	60	6	6	60	60	6	192	60	1920
1	12	1	120	12	12	120	120	12	96	120	960
1	24	1	240	24	24	240	240	24	48	240	480
1	48	1	480	48	48	480	480	48	24	480	240
1	96	1	960	96	96	960	960	96	12	960	120
1	192	1	1920	192	192	1920	1920	192	6	1920	60
1	384	1	3840	384	384	3840	3840	384	3	3840	30
1	768	1	7680	768	768	7680	7680	768	1.5	7680	15

* Additional patterns exist, and can be analyzed as appropriate.

One analysis, for the six sample sizes indicated above and with one repetition, was performed. This analysis included the covariance values calculated using the SDx and SDy values indicated in Pattern One. The covariance matrix and results are summarized in Tables 38 and 39 below. The output from the Lavaan SEM is provided in Appendix C.

Table 38

Covariance Matrix

Model Fit Index	z	X1	X2	X3
z	4.8			
X1	0.6	1.2		
X2	38.4	0.3	76.8	
X3	2.4	19.2	0.15	9.6

Table 39

Model Fit Indices for Correlation Value of 0.1

Sample Size	50	100	150	200	300	500
Chi-Squared P Value	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
RMSEA Lower	0.0	0.0	0.0	0.0	0.0	0.0
RMSEA Upper	0.0	0.0	0.0	0.0	0.0	0.0
SRMR	**1.9**	**1.9**	**1.9**	**1.9**	**1.9**	**1.9**
CFI	1.0	1.0	1.0	1.0	1.0	1.0

Note: Bold text indication of poor model fit.

The model fit results from analysis were mixed. The Chi-Squared and SRMR test results were an indication of a poor model fit. The RMSEA Lower, RMSEA Upper and CFI test results were an indication of a good model fit. The model fit index values did not vary significantly between sample sizes.

As the covariance matrix was compiled using patterned values of random formulation, an indication of appropriate model fit test results would be a poor model fit. It would be interesting to see the results for smaller correlation values, such as 0.025 and 0.05, to determine the lowest correlation values the model fit index results for the Chi-Squared and CFI tests resulted in a poor model fit.

APPENDIX A

Table A1

Upper Tail Probabilities of the Chi-Squared Distribution (Neave & Worthington, 1988)

	α			α	
ν	5%	1%	ν	5%	1%
1	3.841	6.635	26	38.885	45.642
2	5.991	9.210	27	40.113	46.963
3	7.815	11.345	28	41.337	48.278
4	9.488	13.277	29	42.557	49.588
5	11.070	15.086	30	43.773	50.892
6	12.592	16.812	31	44.985	52.191
7	14.067	18.475	32	46.194	53.486
8	15.507	20.090	33	47.400	54.776
9	16.919	21.666	34	48.602	56.061
10	18.307	23.209	35	49.802	57.342
11	19.675	24.725	36	50.998	58.619
12	21.026	26.217	37	52.192	59.893
13	22.362	27.688	38	53.384	61.162
14	23.685	29.141	39	54.572	62.428
15	24.996	30.578	40	55.758	63.691
16	26.296	32.000	45	61.656	69.957
17	27.587	33.409	50	67.505	76.154
18	28.869	34.805	60	79.082	88.379
19	30.144	36.191	70	90.531	110.43
20	31.410	37.566	80	101.88	112.33
21	32.671	38.932	90	113.15	124.12
22	33.924	40.289	100	124.34	135.81
23	35.172	41.638			
24	36.415	42.980			
25	37.652	44.314			

APPENDIX B

sample size=250
repetitions=10000
mean correlation=0.463424754334544
mean DOF=3
chisq a<0.05=1-> This means that p-chi-squared was less than 0.05 5% of the time
chisq a<0.01=1-> This means that p-chi-squared was less than 0.01 1% of the time
RMSEA (lower)<=0.05=2e-04-> This means that RMSEA Lower is less than 0.05 0% of the time
RMSEA (lower)>0.05=0.9998-> This means that RMSEA Lower is greater than 0.05 100% of the time
RMSEA (upper)<=0.1=0-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)>0.1=1-> This means that RMSEA Upper is greater than 0.1 100% of the time
SRMR<=0.09=5e-04-> This means that SRMR is less than 0.09 0% of the time
SRMR>0.09=0.9995-> This means that SRMR is greater than 0.09 100% of the time
CFI<0.9=0.9068-> This means that CFI is less than 0.9 91% of the time
CFI<0.75=0.4115-> This means that CFI is less than 0.75 41% of the time
CFI<0.5=0.0295-> This means that CFI is less than 0.5 3% of the time

Figure B1. Lavaan Output for Sample Size of 250 and Four Variables, Repetitions = 10,000.

sample size=100
repetitions=10000
mean correlation=0.433938716916882
mean DOF=3
chisq a<0.05=1-> This means that p-chi-squared was less than 0.05 100% of the time
chisq a<0.01=1-> This means that p-chi-squared was less than 0.01 100% of the time
RMSEA (lower)<=0.05=0.0108-> This means that RMSEA Lower is less than 0.05 1% of the time
RMSEA (lower)>0.05=0.9892-> This means that RMSEA Lower is greater than 0.05 99% of the time
RMSEA (upper)<=0.1=0-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)>0.1=1-> This means that RMSEA Upper is greater than 0.1 100% of the time
SRMR<=0.09=0.0028-> This means that SRMR is less than 0.09 0% of the time
SRMR>0.09=0.9972-> This means that SRMR is greater than 0.09 100% of the time
CFI<0.9=0.8297-> This means that CFI is less than 0.9 83% of the time
CFI<0.75=0.4133-> This means that CFI is less than 0.75 41% of the time
CFI<0.5=0.1526-> This means that CFI is less than 0.5 15% of the time

Figure B2. Lavaan Output for Sample Size of 100 and Four Variables, Repetitions = 10,000.

sample size=50
repetitions=10000
mean correlation=0.412394902961997
mean DOF=3
chisq a<0.05=1-> This means that p-chi-squared was less than 0.05 100% of the time
chisq a<0.01=1-> This means that p-chi-squared was less than 0.01 100% of the time
RMSEA (lower)<=0.05=0.0539-> This means that RMSEA Lower is less than 0.05 5% of the time
RMSEA (lower)>0.05=0.9461-> This means that RMSEA Lower is greater than 0.05 10% of the time
RMSEA (upper)<=0.1=3e-04-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)>0.1=0.9997-> This means that RMSEA Upper is greater than 0.1 100% of the time
SRMR<=0.09=0.0055-> This means that SRMR is less than 0.09 1% of the time
SRMR>0.09=0.9945-> This means that SRMR is greater than 0.09 99% of the time
CFI<0.9=0.8058-> This means that CFI is less than 0.9 81% of the time
CFI<0.75=0.4441-> This means that CFI is less than 0.75 44% of the time
CFI<0.5=0.2316-> This means that CFI is less than 0.5 23% of the time

Figure B3. Lavaan Output for Sample Size of 50 and Four Variables, Repetitions = 10,000.

sample size=30
repetitions=10000
mean correlation=0.397398624898292
mean DOF=3
chisq a<0.05=1-> This means that p-chi-squared was less than 0.05 100% of the time
chisq a<0.01=1-> This means that p-chi-squared was less than 0.01 100% of the time
RMSEA (lower)<=0.05=0.0913-> This means that RMSEA Lower is less than 0.05 9% of the time
RMSEA (lower)>0.05=0.9087-> This means that RMSEA Lower is greater than 0.05 91% of the time
RMSEA (upper)<=0.1=0.0012-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)>0.1=0.9988-> This means that RMSEA Upper is greater than 0.1 100% of the time
SRMR<=0.09=0.0053-> This means that SRMR is less than 0.09 1% of the time
SRMR>0.09=0.9947-> This means that SRMR is greater than 0.09 99% of the time
CFI<0.9=0.8252-> This means that CFI is less than 0.9 83% of the time
CFI<0.75=0.5033-> This means that CFI is less than 0.75 50% of the time
CFI<0.5=0.2795-> This means that CFI is less than 0.5 30% of the time

Figure B4. Lavaan Output for Sample Size of 30 and Four Variables, Repetitions = 10,000.

sample size=20
repetitions=10000
mean correlation=0.386482079166151
mean DOF=3
chisq a<0.05=1-> This means that p-chi-squared was less than 0.05 100% of the time
chisq a<0.01=1-> This means that p-chi-squared was less than 0.01 100% of the time
RMSEA (lower)<=0.05=0.1224-> This means that RMSEA Lower is less than 0.05 12% of the time
RMSEA (lower)>0.05=0.8776-> This means that RMSEA Lower is greater than 0.05 88% of the time
RMSEA (upper)<=0.1=0.0011-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)>0.1=0.9989-> This means that RMSEA Upper is greater than 0.1 100% of the time
SRMR<=0.09=0.0044-> This means that SRMR is less than 0.09 0% of the time
SRMR>0.09=0.9956-> This means that SRMR is greater than 0.09 100% of the time
CFI<0.9=0.8387-> This means that CFI is less than 0.9 84% of the time
CFI<0.75=0.5637-> This means that CFI is less than 0.75 56% of the time
CFI<0.5=0.311-> This means that CFI is less than 0.5 31% of the time

Figure B5. Lavaan Output for Sample Size of 20 and Four Variables, Repetitions = 10,000.

sample size=10
repetitions=10000
mean correlation=0.0421997270262573
mean DOF*=0.3432
chisq a**<0.05=0.1144-> This means that p-chi-squared was less than 0.05 11% of the time
chisq a**<0.01=0.1143-> This means that p-chi-squared was less than 0.01 11% of the time
RMSEA (lower)**<=0.05=0.0186-> This means that RMSEA Lower is less than 0.05 2% of the time
RMSEA (lower)**>0.05=0.0958-> This means that RMSEA Lower is greater than 0.05 10% of the time
RMSEA (upper)**<=0.1=0-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)**>0.1=0.1144-> This means that RMSEA Upper is greater than 0.1 11% of the time
SRMR**<=0.09=0-> This means that SRMR is less than 0.09 0% of the time
SRMR**>0.09=0.1144-> This means that SRMR is greater than 0.09 11% of the time
CFI**<0.9=0.0994-> This means that CFI is less than 0.9 10% of the time
CFI**<0.75=0.0766-> This means that CFI is less than 0.75 8% of the time
CFI**<0.5=0.0413-> This means that CFI is less than 0.5 4% of the time

---

*Mean degrees of freedom did not result in a whole number.

**Results indicated that RMSEA Lower was less than 0.05 eleven percent of the time and greater than 0.05 for ten percent of the repetitions. These values do not add up to 100%.

Figure B6. Lavaan Output for Sample Size of 10 and Four Variables, Repetitions = 10,000.

sample size=450
repetitions=10000
mean correlation x1&z=0.000462070478175039
mean correlation x2&z=0.0025020921013811
mean correlation x3&z=0.00227332378855168
mean correlation x4&z=0.000598797833551206
mean DOF= numeric(0)*
chisq a**<0.05=0.0325-> This means that p-chi-squared was less than 0.05 3% of the time
chisq a**<0.01=0.0325-> This means that p-chi-squared was less than 0.01 3% of the time
RMSEA (lower)**<=0.05=0-> This means that RMSEA Lower is less than 0.05 0% of the time
RMSEA (lower)**>0.05=0.0325-> This means that RMSEA Lower is greater than 0.05 3% of the time
RMSEA (upper)**<=0.1=0-> This means that RMSEA Upper is less than 0.1 0% of the time
RMSEA (upper)**>0.1=0.0325-> This means that RMSEA Upper is greater than 0.1 3% of the time
SRMR**<=0.09=0-> This means that SRMR is less than 0.09 0% of the time
SRMR**>0.09=0.0325-> This means that SRMR is greater than 0.09 3% of the time
CFI**<0.9=0.0325-> This means that CFI is less than 0.9 3% of the time
CFI**<0.75=0.0325-> This means that CFI is less than 0.75 3% of the time
CFI**<0.5=0.0325-> This means that CFI is less than 0.5 3% of the time

---

*RStudio was not able to calculate the degrees of freedom.

** Results indicated that RMSEA Lower was less than 0.05 zero percent of the time and greater than 0.05 for three percent of the repetitions. These values do not add up to 100%.

Figure B7. Lavaan Output for Sample Size of 250 and Five Variables, Repetitions = 10,000.

APPENDIX C

```
lavaan (0.5-18) converged normally after   8 iterations

  Number of observations                          50
  Estimator                                     ML
  Minimum Function Test Statistic              0.000
  Degrees of freedom                              3
  P-value (Chi-square)                        1.000

  Comparative Fit Index (CFI)                  1.000
  Tucker-Lewis Index (TLI)                     1.003

  Loglikelihood user model (H0)              617.304
  Loglikelihood unrestricted model (H1)      617.304

  RMSEA                                        0.000
  90 Percent Confidence Interval          0.000  0.000
  P-value RMSEA <= 0.05                        1.000

  SRMR                                         1.877

                   Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  z ~
    x1               -0.199    0.043   -4.614    0.000
    x2                0.499    0.006   85.754    0.000
    x3                0.264    0.016   16.025    0.000

Covariances:
  x1 ~~
    x2                0.000
    x3                0.000
  x2 ~~
    x3                0.000

Variances:
    z                 0.073    0.015
    x1                0.783    0.157
    x2               42.946    8.589
    x3                5.368    1.074
```

Figure C1. Lavaan Output for Sample Size of 50 and Four Variables, Correlation = 0.1.

```
lavaan (0.5-18) converged normally after   8 iterations

  Number of observations                         100
  Estimator                                 ML
  Minimum Function Test Statistic              0.000
  Degrees of freedom                             3
  P-value (Chi-square)                        1.000

  Comparative Fit Index (CFI)                 1.000
  Tucker-Lewis Index (TLI)                    1.001

  Loglikelihood user model (H0)            1234.607
  Loglikelihood unrestricted model (H1)      1234.607

  RMSEA                                  0.000
  90 Percent Confidence Interval          0.000  0.000
  P-value RMSEA <= 0.05                        1.000

  SRMR                                   1.877

                 Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  z ~
    x1             -0.199    0.031   -6.505    0.000
    x2              0.499    0.004  121.013    0.000
    x3              0.264    0.012   22.624    0.000

Covariances:
  x1 ~~
    x2              0.000
    x3              0.000
  x2 ~~
    x3              0.000

Variances:
    z              0.074    0.010
    x1              0.788    0.111
    x2             43.386    6.136
    x3              5.423    0.767
```

Figure C2. Lavaan Output for Sample Size of 100 and Four Variables, Correlation = 0.1.

```
lavaan (0.5-18) converged normally after   8 iterations

  Number of observations                          150
  Estimator                                    ML
  Minimum Function Test Statistic                0.000
  Degrees of freedom                            3
  P-value (Chi-square)                       1.000

  Comparative Fit Index (CFI)                 1.000
  Tucker-Lewis Index (TLI)                  1.001

  Loglikelihood user model (H0)            1851.911
  Loglikelihood unrestricted model (H1)      1851.911

  RMSEA                                    0.000
  90 Percent Confidence Interval         0.000  0.000
  P-value RMSEA <= 0.05                         1.000

  SRMR                                    1.877

                 Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  z ~
    x1             -0.199    0.025   -7.959    0.000
    x2              0.499    0.003  148.106    0.000
    x3              0.264    0.010   27.693    0.000

Covariances:
  x1 ~~
    x2              0.000
    x3              0.000
  x2 ~~
    x3              0.000

Variances:
    z              0.074    0.009
    x1              0.790    0.091
    x2             43.533    5.027
    x3              5.442    0.628
```

Figure C3. Lavaan Output for Sample Size of 150 and Four Variables, Correlation = 0.1.

```
lavaan (0.5-18) converged normally after   8 iterations

  Number of observations                          200
  Estimator                                     ML
  Minimum Function Test Statistic                0.000
  Degrees of freedom                              3
  P-value (Chi-square)                         1.000

  Comparative Fit Index (CFI)                   1.000
  Tucker-Lewis Index (TLI)                     1.001

  Loglikelihood user model (H0)             2469.215
  Loglikelihood unrestricted model (H1)       2469.215

  RMSEA                                    0.000
  90 Percent Confidence Interval         0.000  0.000
  P-value RMSEA <= 0.05                         1.000

  SRMR                                    1.877

                Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  z ~
    x1            -0.199    0.022   -9.186    0.000
    x2             0.499    0.003  170.958    0.000
    x3             0.264    0.008   31.969    0.000

Covariances:
  x1 ~~
    x2             0.000
    x3             0.000
  x2 ~~
    x3             0.000

Variances:
    z              0.074    0.007
    x1             0.791    0.079
    x2            43.607    4.361
    x3             5.451    0.545
```

Figure C4. Lavaan Output for Sample Size of 200 and Four Variables, Correlation = 0.1.

```
lavaan (0.5-18) converged normally after   8 iterations

  Number of observations                          300
  Estimator                                     ML
  Minimum Function Test Statistic                 0.000
  Degrees of freedom                              3
  P-value (Chi-square)                         1.000

Model test baseline model:

  Comparative Fit Index (CFI)                   1.000
  Tucker-Lewis Index (TLI)                     1.000

  Loglikelihood user model (H0)             3703.822
  Loglikelihood unrestricted model (H1)       3703.822

  RMSEA                                     0.000
  90 Percent Confidence Interval         0.000  0.000
  P-value RMSEA <= 0.05                         1.000

  SRMR                                     1.877

                 Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  z ~
    x1            -0.199    0.018  -11.245    0.000
    x2             0.499    0.002  209.307    0.000
    x3             0.264    0.007   39.143    0.000

Covariances:
  x1 ~~
    x2             0.000
    x3             0.000
  x2 ~~
    x3             0.000

Variances:
    z             0.075    0.006
    x1             0.792    0.065
    x2            43.680    3.566
    x3             5.460    0.446
```

Figure C5. Lavaan Output for Sample Size of 300 and Four Variables, Correlation = 0.1.

```
lavaan (0.5-18) converged normally after   8 iterations

  Number of observations                          500
  Estimator                                     ML
  Minimum Function Test Statistic                0.000
  Degrees of freedom                              3
  P-value (Chi-square)                          1.000

  Comparative Fit Index (CFI)                   1.000
  Tucker-Lewis Index (TLI)                      1.000

  Loglikelihood user model (H0)              6173.036
  Loglikelihood unrestricted model (H1)      6173.036

  RMSEA                                         0.000
  90 Percent Confidence Interval          0.000  0.000
  P-value RMSEA <= 0.05                         1.000

  SRMR                                          1.877

                Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  z ~
    x1            -0.199    0.014  -14.511    0.000
    x2             0.499    0.002  270.140    0.000
    x3             0.264    0.005   50.523    0.000

Covariances:
  x1 ~~
    x2             0.000
    x3             0.000
  x2 ~~
    x3             0.000

Variances:
    z              0.075    0.005
    x1             0.793    0.050
    x2            43.739    2.766
    x3             5.467    0.346
```

Figure C6. Lavaan Output for Sample Size of 500 and Four Variables, Correlation = 0.1.

REFERENCES

Alavifar, A., Karimimalayer, M., & Anuar, M.K. (2012). Structural Equation Modeling VS Multiple Regression. *Engineering Science and Technology: An International Journal, 2(2), 326-329.* Retrieved from http://www.estij.org/papers/vol2no22012/25vol2no2.pdf

Arbuckle, J. L. (2011). *IBMor SPSSor AMOSTM 20 User's Guide.* Retrieved from . ftp://public.dhe.ibm.com/software/analytics/spss/documentation/amos/20.0/en/Manuals/IBM_SPSS_Amos_User_Guide.pdf

Bagozzi, R. P. & Youjae, I. (1989). On the Use of Structural Equation Models in Experimental Designs. *Journal of Marketing Research, (26)3, 271-284.* Retrieved from http://www.jstor.org.proxy.lib.wayne.edu/stable/pdf/3172900.pdf

Bollen, K.A. (1989). A New Incremental Fit Index for General Structural Equation Models. *Sociological Methods & Research, 17(3), 303-316.* Retrieved from http://smr.sagepub.com.proxy.lib.wayne.edu/content/17/3/303

Byrne, B.M. (1994). *Structural Equation Modeling with EQS and EWS/Windows: Basic Concepts, Applications, and Programming.* Thousand Oaks, CA: Sage Publications.

Cheong, Y. S., Reynolds, G. K., Salman, A. D., & Hounslow, M. J. (2004). Modelling Fragment Size Distribution Using Two-Parameter Weibull Equation. *International Journal of Mineral Processing, 74, 227-237.* Retrieved from http://ac.els-cdn.com/S0301751604000572/1-s2.0-S0301751604000572-main.pdf?_tid=f846face-0362-11e5-a84e-00000aab0f01&acdnat=1432616111_657d997ad213b7b9ea6b1ccd1ba6e421

Chin, W. W. (1998). Issues and Opinion on Structural Equation Modeling. *MIS Quarterly, (22)1, xii-xvi.* Retrieved from http://link.springer.com.proxy.lib.wayne.edu/article/10.1007%2FBF02295939 http://www.jstor.org.proxy.lib.wayne.edu/stable/pdf/249674.pdf?acceptTC=true

Ewing, R., Hamidi, S., Gallivan, F., Nelson, A. C., & Grace, J. B. (2014). Structural Equation Models of VMT Growth in US Urbanised Areas. *Urban Studies, (51)14, 3079-3096.* Retreived from http://usj.sagepub.com.proxy.lib.wayne.edu/content/51/14/3079.full.pdf+html

Guralnik, J. M., Ferrucci, L., Pieper, C. F., Leveille, S. G., Markides, K. S., Ostire, G. V., Studenski, S., Berkman, L. F., Wallace, R. B. (2000). Lower Extremity Function and Subsequent Disability: Consistency Across Studies, Predictive Models, and Value of Gait Speed Alone Compared With the Short Physical Performance Battery. *Journal of Gerontology, (55A)4, M221-M231.* Retrieved from http://biomedgerontology.oxfordjournals.org/content/55/4/M221.full.pdf+html

Gullen, J.A. (2000). *Goodness of Fit as a Single Factor Structural Equation Model* (Unpublished doctoral dissertation). Wayne State University, Detroit.

Hesketh, S. R., Skrondal, A., & Pickles, A. (2004). Generalized Multilevel Structural Equation Modeling. *Psychometrika, (69)2, 167-190.* Retrieved from http://link.springer.com.proxy.lib.wayne.edu/article/10.1007%2FBF02295939

Hinkle, D.E., Wiersma, W., & Jurs, S. G. (1998). *Applied Statistics for the Behavioral Sciences.* Bostom, MA: Houghton Mifflin Company.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods, 6(1), 53-60.* Retrieved from http://arrow.dit.ie/cgi/viewcontent.cgi?article=1001&context=buschmanart&sei-redir=1&referer=http%3A%2F%2Fscholar.google.com%2Fscholar%3Fq%3Dfit%2Bindices%2Bsem%26hl%3Den%26as_sdt%3D0%26as_vis%3D1%26oi%3Dscholart%26sa%3DX%26ei%3DUVBRVcWyJtOXygTiqIG4Aw%26ved%3D0CBsQgQMwAA#search=%22fit%20indices%20sem%22

Hoyle, R.H. (Ed.). (1995). *Structural Equation Modeling: Concepts, Issues, and Applications.* Thousand Oaks, CA: Sage Publications.

Hu, L. & Bentler, P.M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, (6)1, 1-55.* Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/10705519909540118#.VVlhe-9_mmE

Iacobucci, D. (2009). Structural Equations Modeling: Fit Indices, Sample Size, and Advanced Topics. *Journal of Consumer Psycology, (20), 90-98.* Retrieved from http://www.sciencedirect.com.proxy.lib.wayne.edu/science/article/pii/S1057740809001120

Kelly, S. (2011). Do Homes that are More Energy Efficient Consume Less Energy?: A Structural Equation Model of the English Residential Sector. *Energy, (36)9, 5610-5620.* Retrieved from http://ac.els-cdn.com/S0360544211004579/1-s2.0-S0360544211004579-main.pdf?_tid=1a92a91c-fdd1-11e4-a00c-00000aacb35d&acdnat=1432003706_7af08ef683fc4426590c6888cf8d6151

Kline, R.B. (2011). *Principles and Practice of Structural Equation Modeling* (3rd ed.). New York, NY: The Guilford Press.

Matsueda, R.L. (2011). *Key Advances in the History of Structural Equation Modeling*. Retrieved from https://www.csss.washington.edu/Papers/2012/wp114.pdf

Miles, J., & Shevlin, M. (2007). A Time and a Place for Incremental Fit Indices. *Personality and Individual Differences, (42), 869-874.* Retrieved from http://ac.els-cdn.com/S0191886906003874/1-s2.0-S0191886906003874-main.pdf?_tid=db66ccb6-fcf7-11e4-b6da-00000aacb362&acdnat=1431910399_ee783c2275e8ba7adc0a77aec9f12383

Naresky, J. J. (1970). Reliability Definitions. *IEEE Transactions on Reliability, (19)4, 198-200.* Retrieved from http://ieeexplore.ieee.org.proxy.lib.wayne.edu/stamp/stamp.jsp?tp=&arnumber=5216447

Neave, H.R., & Wrothington, P.L. (1988). *Distribution Free Tests.* Boston, MA: Unwin Hyman Inc.

Nichols, J. F., Morgan, C. G., Chabot, L. E., Sallis, J. F., Calfas, K. J. (2000). Assessment of Physical Activity With the Computer Science and Applications, Inc., Accelerometer: Laboratory Versus Field Validation. *Research Quarterly for Exercise and Sport, (71)1, 36-43.* Retrieved from http://search.proquest.com.proxy.lib.wayne.edu/docview/218495524/fulltextPDF/8F9816224794485CPQ/1?accountid=14925

Nunnally, J.C. & Bernstein, I.H. (1994). *Psychometric Theory* (3rd ed.). New York, NY: McGraw-Hill, Inc.

Pearson, K. (1900). On the Criterion that a Given System of Deviations from the Probable in the Case of a Correlated System of Variables is Such that it Can be Reasonably Supposed to have Arisen from Random Sampling. *Philosophical Magazine Series 5, (50)302, 157-175.* Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/14786440009463897#.VVlJwe9_mmE

Petit, T., Wilt, M., Velten, M., Millon, R., Rodier, J. F., Borel, C., Mors, R., Haegele, P., Eber, M., Ghnassia, J. P. (2004). Comparative Value of Tumour Grade, Hormonal Receptors, Ki-67, HER-2 and Topoisomerase II Alpha Status as Predictive Markers in Breast Cancer Patients Treated with Neoadjuvant Anthracycline-Based Chemotherapy. *European Journal of Cancer, (40), 205-211.* Retrieved from http://www.sciencedirect.com.proxy.lib.wayne.edu/science/article/pii/S0959804903006750#

Reinartz, W. J., Echambadi, R., & Chin, W. W. (2002). Generating Non-normal Data for Simulation of Structural Equation Models Using Mattson's Method. *Multivariate Behavioral Research, 37(2), 227-244.* Retrieved from http://www.tandfonline.com.proxy.lib.wayne.edu/doi/pdf/10.1207/S15327906MBR3702_03

Schermelleh-Engel, K., Moosbrugger, H., & Muller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research, (8)2, 23-74.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.509.4258&rep=rep1&type=pdf

Sinfield, J. V., Einstein, H. H. (1998). Tunnel Construction Costs for Tube Transportation Systems. *Journal of Construction Engineering and Management, (124), 48-57.* Retrieved from http://ascelibrary.org.proxy.lib.wayne.edu/doi/pdf/10.1061/%28ASCE%290733-9364%281998%29124%3A1%2848%29

Tomarken, A. J. & Waller, N. G. (2005). Structural Equation Modeling: Strengths, Limitations, and Misconceptions. *Annual Review of Clinical Psychology, (1)1, 31-65.* Retrieved from http://www.annualreviews.org.proxy.lib.wayne.edu/doi/pdf/10.1146/annurev.clinpsy.1.102803.144239

Tylka, L. T. & Subich, L. M. (2004). Examining a Multidimensional Model of Eating Disorder Symptomatology Among College Women. *Journal of Psychology, (51)3, 314-328.* Retrieved from http://up7af9tu5s.search.serialssolutions.com/?ctx_ver=Z39.88-2004&ctx_enc=info%3Aofi%2Fenc%3AUTF-8&rfr_id=info:sid/summon.serialssolutions.com&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&rft.genre=article&rft.atitle=Examining+a+Multidimensional+Model+of+Eating+Disorder+Symptomatology+Among+College+Women&rft.jtitle=Journal+of+Counseling+Psychology&rft.au=Tylka%2C+Tracy+L&rft.au=Subich%2C+Linda+Mezydlo&rft.date=2004&rft.issn=0022-0167&rft.eissn=1939-2168&rft.volume=51&rft.issue=3&rft.spage=314&rft.epage=328&rft_id=info:doi/10.1037%2F0022-0167.51.3.314&rft.externalDBID=n%2Fa&rft.externalDocID=10_1037_0022_0167_51_3_314&paramdict=en-US

Yeh, S. F., Lin, M. D., & Tsai, K. T. (2008). Development of Cost Functions for Open-Cut and Jacking Methods for Sanitary Sewer System Construction in Central Taiwan. *Practice Periodical of Hazardous, Toxic, and Radioactive Waste Management, (12), 282-289.* Retrieved from http://ascelibrary.org.proxy.lib.wayne.edu/doi/pdf/10.1061/%28ASCE%291090-025X%282008%2912%3A4%28282%29

Weston, R. & Gore, P.A. (2006). A Brief Guide to Structural Equation Modeling. *The Counseling Psychologist, 34(5), 719-751.* Retrieved from http://ascelibrary.org.proxy.lib.wayne.edu/doi/pdf/10.1061/%28ASCE%291090-025X%282008%2912%3A4%28282%29 http://tcp.sagepub.com.proxy.lib.wayne.edu/content/34/5/719.full.pdf+html

Wright, S. (1918). On the Nature of Size Factors. *Genetics, 3(1), 367-374.* Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1200442/pdf/367.pdf

ABSTRACT

ESTIMATING EFFECTS OF NON-NORMALITY IN ASSESSING STRUCTURAL EQUATION MODEL FIT FOR USE OF PHYSICAL SCIENCE DATA

by

SARAH ALTA ROSE

May 2016

Advisor: Dr. Barry Markman

Major: Education (Evaluation and Research)

Degree: Doctor of Philosophy

The purpose of this study was to evaluate the sensitivity of selected fit index statistics in determining model fit when the distribution varied from normality, as is typically true of data research for the physical sciences. SEM is a popular statistical method and is used in many physical and social behavioral science research projects; however, the sensitivity of the model fit indices when normality is violated had never been estimated.

The original intent for performing the research was to analyze the legitimacy of the model fit indices' results against three different types of distributions. One data distribution contained five variables (Torque at Transmission, Engine Speed, Vehicle Speed, Accelerator Pedal Position, and Fuel Used). The variables were assessed using an SEM, and a Monte Carlo simulation of 10,000 iterations for varying sample sizes of n = 10, 20, 30, 50, 100, and 250. It was determined that an indication of poor model fit occurred with greater consistency as the sample size of the data set increased. It was also determined that the magnitude of the correlation decreased as

sample size decreased, and as the correlation approached zero the model fit resulted in illogical results.

Additional Monte Carlo simulations were therefore conducted, with 1,000 repetitions, and varying magnitudes of correlation matrix values randomly selected from a range of a base value plus or minus 0.015. Twenty four Monte Carlo simulations were performed, with the base value increasing from 0.04 to 0.27. As the correlation matrix values were increased in magnitude, the results of the model fit indices became first illogical and then finally logical, with an increasing indication of a poor model fit. At a certain specified magnitude, the results of the model fit indices were an indication of a poor model fit for the model fit index studied for all Monte Carlo repetitions.

These results were forwarded to a subject matter expert. This expert, in citing Kline (2011), opined the most likely explanation for the study results was the default method of Maximum Likelihood that assumes variables are unstandardized. When variables are standardized, the results could be incorrect. The subject matter expert concluded that no systematic Monte Carlo study could be conducted by inputting an incrementally increasing correlation matrix, such as was attempted in this study.

AUTOBIOGRAPHICAL STATEMENT

Sarah Alta Rose received her Bachelor of Science (Magna Cum Laude) in Architecture and Civil Engineering in 2007 and her Master of Science in Civil Engineering in 2009 from Lawrence Technological University in Southfield, Michigan. Specializing in Structural Engineering, Sarah utilized her academic knowledge working for a consulting company on underground design construction projects, including one record breaking sinking shaft project that won a 2006 Clean Water State Revolving Fund PISCES Award. While working on this ground breaking project, Sarah developed an interest for research and data analytics resulting in the pursuit of her Doctor of Philosophy in Evaluation and Research (EER) at Wayne State University. Currently, Sarah works for Ford Motor Company as a data analyst, working on new research projects for upcoming vehicle models. Sarah also teaches as adjunct/part time faculty Wayne State University.