

5-1-2010

Impact of Measurement Model Modification on Structural Parameter Integrity When Measurement Model is Misspecified

Weihua Fan

University of Houston, wfan2@uh.edu

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Fan, Weihua (2010) "Impact of Measurement Model Modification on Structural Parameter Integrity When Measurement Model is Misspecified," *Journal of Modern Applied Statistical Methods*: Vol. 9 : Iss. 1 , Article 12.
DOI: 10.22237/jmasm/1272687060

Impact of Measurement Model Modification on Structural Parameter Integrity When Measurement Model is Misspecified

Weihua Fan
University of Houston

In the process of model modification, parameters of residual covariances are often treated as free parameters to improve model fit. However, the effect of such measurement model modifications on the important structural parameter estimates under various measurement model misspecifications has not been systematically studied. Monte Carlo simulation was conducted to compare structural estimates before and after measurement model modifications of adding residual covariances under varying sample sizes and model misspecifications. Results showed that researchers should pay attention when such measurement model modifications are made to initially misspecified model with missing path(s).

Key words: Structural equation modeling; modification indices; Lagrange multiplier (LM) tests; residual covariance; misspecification.

Introduction

Model modification, also known as specification search, has been widely used in the application of structural equation modeling, in hope of improving models. After obtaining a model that fails to meet accepted goodness of fit standards, many researchers frequently turn toward model modification information in an attempt to find a parsimonious model to fit the sample data. Two approaches commonly used for model modification are the Lagrange Multiplier test (LM test; also referred to as modification index) and the Wald (W) test. The LM test reduces constraints by freely estimating parameters such as residual covariances that are currently fixed (usually to zero). The W test increases constraints by fixing parameters that are currently free (Bentler, 1995; Chou & Bentler, 1990). The more applied approach is usually the LM test, because freely estimating parameters rather than fixing parameters improves model fit.

Anderson and Gerbing (1988) proposed a two-step SEM process as some guidance for

latent variable path models which is commonly recommended in practice (see, Kline, 2004; Schumacker & Lomax, 2004). The SEM process consists of two steps: the measurement step and the structural step. The measurement step in the process considers the measurement model, which specifies the relations between the underlying factors and the measured variables. It allows researchers opportunities to improve the data-model fit through model modification within the measurement model while temporarily inserting a saturated latent structure. After obtaining satisfactory data-model fit in the first step, the second step involves the structural model which hypothesizes relations between the latent variables. Assessing the structural relations in an SEM application is usually the focal point of an investigation. Examples of this two-step process include studies by Mattanah, Hancock, and Brand (2004), Joiner, Leveson, and Langfield-Smith (2002) and Chong and Chong (2002). Also see Mulaik and Millsap (2000) for a four-step process and Green, Thompson and Poirier (1999) for a 2-stage specification search procedure, and Green, Thompson and Poirier (2001) for an adjusted Bonferroni method for eliminating parameters in specification searches.

Although the application of the model modification procedures is commonly observed

Weihua Fan is an Assistant Professor in the College of Education, Department of Educational Psychology. Email: wfan2@uh.edu.

MEASUREMENT MODEL MODIFICATION UNDER MISSPECIFICATION

across disciplines (Brekler, 1990; Hutchinson, 1998; MacCallum, Roznowski & Necowitz, 1992), it is data-driven in nature and is characterized by capitalization on chance (e.g., MacCallum, Roznowski & Necowitz, 1992). Researchers have thus become prudent with model modifications regarding important theory related parameters such as structural parameters or loading parameters. In the meantime, parameters of residual covariances, which are much less theoretically concerned, are often treated as free parameters to help improve the model-data fit. This is exemplified in work by Newcomb and Bentler (1988) in which 77 residual covariances were added prior to structural analyses.

Given that the primary purpose of latent variable path models is to assess theoretical relations between latent variables (e.g., Kerlinger, 1986), it is not uncommon that parameters of residual covariances are used in abundance to help improve model fit. The liberal manner in which residual covariances are made to measurement models does not appear to be due to a prevailing belief that the true measurement model will be found as a result; research (see, e.g., MacCallum, Roznowski & Necowitz, 1992) has documented that sample-based respecifications seldom arrive at the true population model with consistency.

Two types of errors can result in model modifications such as incorrectly reducing fixed constraints: (1) errors due to sampling fluctuation, and (2) errors due to misspecification (Green, Thompson & Poirier, 1999). Errors related to sampling fluctuation occur if model modifications fit the specific characteristics of sample data but not the population. Errors due to misspecification can occur in two situations. When the initially hypothesized model is correctly specified, model modification of freely estimated correctly fixed parameters to maximally increase model fit is unnecessary and results in Type I error capitalizations on random sample covariation. Fan and Hancock (2006) have shown that the overspecification of measurement model modifications impacts the structural parameter estimates under certain conditions, however, this effect is usually negligible. When the initially hypothesized model is incorrectly specified,

which is almost inevitable in practice, model modifications to revise the model by freely estimating additional parameters can lead to two possible results: a revised model with less misspecification errors than the initially hypothesized model or a revised model with a greater number of errors due to misspecification. Two common types of measurement model misspecification include: (1) relevant parameters are incorrectly fixed to zero, and (2) irrelevant parameters that should be fixed to zero are freely estimated.

If a researcher starts with an initial model with incorrectly fixed relevant parameters set to zero, it is expected that freely estimating incorrectly fixed parameters such as cross-loadings or residual covariances would be beneficial to the structural parameters. However, incorrectly fixing relevant parameters to zero can manifest in significant residual covariances, which exist as a function of fixing the truly non-zero parameters (e.g., cross-loadings) to zero. With this scenario it is likely that the LM test also suggests to freely estimate residual covariances that are correctly fixed to zero, this results in a model with more misspecification errors. In addition, if a researcher starts with an initial model incorrectly freely estimating irrelevant parameters that should be fixed to zero, it is expected that the LM test suggesting more freely estimating incorrectly fixed parameters will lead to a model with more misspecification errors. This begs the question of whether the prevailing modification of freely estimating residual covariances is indeed price-free to our structural parameter estimates under measurement model misspecifications. No study of specification searches has considered this issue, thus the current focus of this investigation.

Recently with the development of the statistical software, a handful of Monte Carlo studies have been conducted to assess the performances of model modifications. Three such studies on model modifications are Hutchinson (1998), Chou and Bentler (2002) and Fan and Hancock (2006). Hutchinson (1998) extended the work of MacCallum, Roznowski and Necowitz (1992) to examine the sampling stability of post hoc model modifications. The study found that modifications tended to be inconsistent unless the sample size is very large

or the model is large. Chou and Bentler (2002) focused on the W test procedure and found a satisfactory success rate of the model modification of structural relationships among factors. However, the results of the test are based on the assumptions of a correct measurement model and a known sequence of latent factors, which pose great challenges to applied researchers. Other relevant studies include Green and Babyak (1997), Green, Thompson and Babyak (1998), Hutchinson (1993), Kaplan (1988, 1989), and Silvia and MacCallum (1988).

Fan and Hancock (2006) investigated the impact of measurement model respecification on structural parameter integrity. Their study compared interfactor correlations before and after measurement model respecifications of crossloadings, intrafactor residual covariances and interfactor residual covariances for a five-factor confirmatory model. The research suggested that some effect on structural parameter estimates arises under conditions of modification in the measurement model; however, in general, the impact is negligible. Although some comfort has been provided for researchers regarding the two-step process of measurement model modification, the study is limited in two ways. First, the study focused on the impact of overspecification of measurement model modifications on structural parameters where paths were added to a correctly specified model. Second, it studied models with interfactor correlations, which appear to be less representing in practice than models involving direct structural relations.

This investigation extends the study of Fan and Hancock (2006) by assessing the effects of freely estimating fixed residual covariances under conditions of measurement model misspecification for a theoretical model involving direct structural relations. This study aims to determine whether model modifications of freely estimating fixed residual covariances are structurally benign under the situations of measurement model misspecification with the goal of gaining insights into the extent to which caution must be exercised in measurement model modification prior to structural model evaluation. For a three-factor confirmatory factor model, Monte Carlo simulation is used to

compare structural estimates before and after measurement model modifications under varying sample sizes and model misspecifications.

Methodology

Model Specification

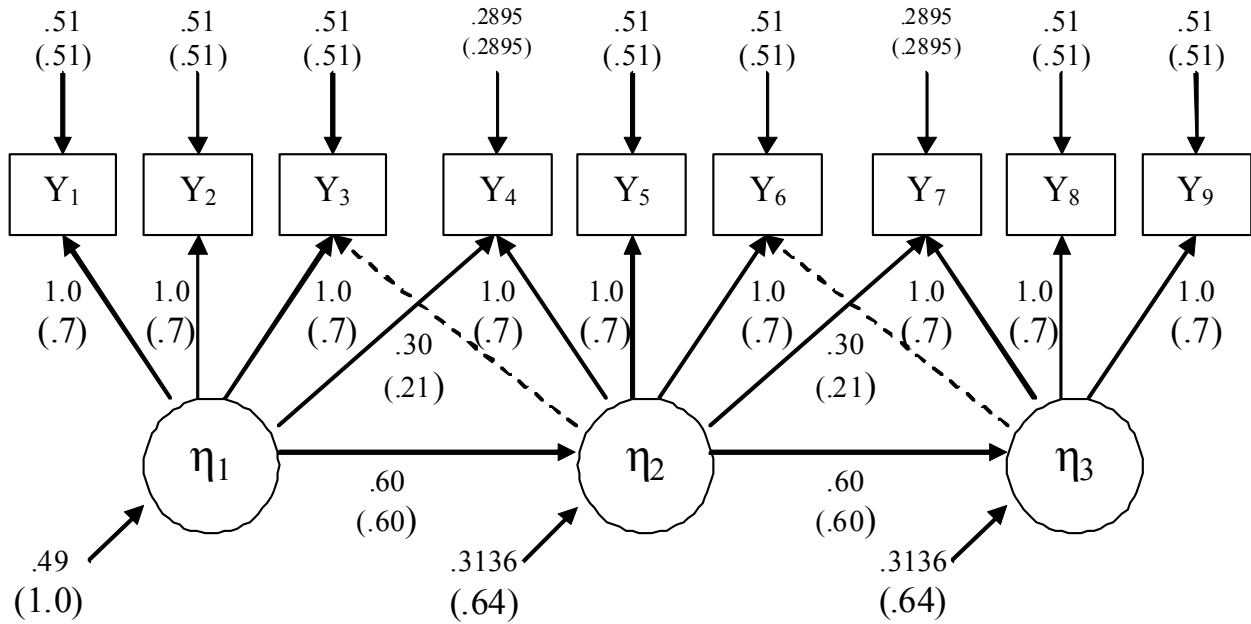
The true model derived from an example illustrated by Paxton, et al. (2001) is a three-factor model measured by nine observed variables. Seven out of nine observed variables load on a single factor and the remaining two load on two factors (See Figure 1). In addition, factor two is regressed on factor one with the coefficient 0.6, and factor three is regressed on factor two with the same coefficient. The model has factor loadings (λ) with an unstandardized value of 1.0 and standardized value of 0.70, while the two cross-loadings have an unstandardized coefficient of 0.30 and a standardized coefficient of 0.21.

Five model specifications are considered in the study. The first model correctly specifies the structure that exists in the population, which is the true model (specification 1). The second model is misspecified by omitting one relevant cross-loading path linking factor 1 with item 4 (specification 2); that is, one relevant cross-loading path is incorrectly fixed to zero. The third model is misspecified by additionally omitting one more relevant cross-loading path linking factor 2 with item 7 (specification 3); that is, one more relevant cross-loading path is incorrectly fixed to zero. The fourth model is misspecified by containing one irrelevant loading from factor 2 to item 3 (specification 4), thus one irrelevant loading path that should be fixed zero is freely estimated. The fifth model is misspecified by containing one additional irrelevant loading from factor 3 to item 6 (specification 5), so one more irrelevant loading path which should be fixed zero is freely estimated.

Note that all modifications of residue covariance suggested to the above models do not exist in the population measurement model. For specification 1, the overspecification of the measurement model simply constitutes Type I error capitalizations on random sample covariation. For the remainder of the specified

MEASUREMENT MODEL MODIFICATION UNDER MISSPECIFICATION

Figure 1: Simulation Model



true models, the modifications of residue covariance are incorrectly freed to maximally improve data-model fit, leading to a revised model with more misspecification errors.

Sample Size, Replications, Data Generation and Modeling

Four different sample sizes are manipulated in the study, ranging from small to large: 100, 200, 400 and 800. A total of 5 (model specifications) \times 4 (sample sizes) = 20 experimental conditions. For each condition, seven separate runs were conducted. Data generation and estimation were carried out using EQS 6.1 (Bentler, 1998) and SAS IML (1990). Within each of the 20 conditions, multivariate normal data were generated and modeled with ML estimation as described by Paxton, et al. (2001).

For each replication within each cell, each of the specified models was first imposed upon the sample data yielding structural estimates of ϕ_1 and ϕ_2 (as well as of loadings and error variances) before any model modification. Second, each of the specified

models with saturated structural correlations was imposed upon the sample data yielding residual covariances suggested by the multivariate Lagrange Multiplier (LM) test. This step is to mimic the measurement step in the two-step modeling process (Anderson & Gerbing, 1988). Next, the residual covariances suggested by the multivariate LM test with cumulative statistical significance $p < .05$ were added to the measurement model and the new structural estimates of ϕ_1 and ϕ_2 were obtained after the model modifications.

In this investigation, the multivariate LM test is restricted to suggesting residual covariances only, the purpose being to diagnose the impact of the most commonly applied measurement model modification on structural relation assessment. Three different measurement model respecifications were examined by the study separately: residual covariances within factors (intrafactor residual covariances) and residual covariances between factors (interfactor residual covariances) and the combination of both. That is, the multivariate LM test is restricted to suggesting only

intrafactor residual covariances in one respecification of each initial model, to suggesting only interfactor residual covariances in another respecification of each initial model, and to suggesting both intrafactor and interfactor residual covariances in yet another modification of each initial model. Thus, each initial solution was subjected to three separate modifications. As the purpose of this study is to diagnose the potential impacts of different types of the residual covariances on structural parameters, different modifications were made in isolation.

To summarize, in the 20 cells of the design, each replication's data were analyzed seven times: (1) initially specified model without any modification, (2) each of the specified models with saturated structural correlations yielding intrafactor residual covariances, (3) specified model with saturated structural correlations yielding interfactor residual covariances, (4) specified models with saturated structural correlations yielding both intrafactor and interfactor residual covariances, (5) respecified model with suggested intrafactor residual covariances added, (6) respecified model with suggested interfactor residual covariances added, and (7) respecified model with both intrafactor and interfactor residual covariances added.

Convergent replications were generated for each condition in which a convergent replication reached convergence within 50 iterations in all analyses based on the same data set and did so without yielding any offending estimates (e.g., Heywood cases). This strategy was adopted because improper solutions in SEM might affect the estimation of structural parameters, thus threatening the results of the study. To maintain 1,400 replications per condition, an initial set of 1,500 replications was generated for each of the conditions.

Analyses

All analyses are based on the 1,400 convergent replications. For each replication on the same sample, the two structural parameters of interest were estimated (a) for the initial model before any modification, (b) after adding to the initial model any intrafactor residual covariances suggested by the multivariate LM test ($p < .05$), (c) after adding to the initial model

any interfactor residual covariances suggested by the multivariate LM test ($p < .05$), and (d) after adding to the initial model any intrafactor and interfactor residual covariances suggested by the multivariate LM test ($p < .05$). As a result, each given cell's 1,400 replications had four estimates for each of the two structural parameters.

After initial examinations of general convergence, three primary analyses were conducted on the 1,400 sets of four structural parameter estimates. First, relative deviation values between parameter estimates before and after model modifications were computed for each cell, each of which indicates the bias of the parameter estimates from the true parameter values. Specifically, for each replication, the deviation was computed for each structural parameter estimate comparing the estimates before and after modification using the formula:

$$\text{Relative bias (RB)} = \hat{\phi}_{\text{after}} - \hat{\phi}_{\text{before}} \quad (1)$$

Thus, for each parameter estimate within a replication, three such deviations were computed, comparing the parameter estimate after each of the three modifications back to the initial estimate. These values are treated as relative biases due to measurement model modifications and are equivalent to the difference in bias with respect to the true parameter value: $(\hat{\phi}_{\text{after}} - \phi) - (\hat{\phi}_{\text{before}} - \phi)$.

The difference between the estimated structural parameter $\hat{\phi}$ and the true parameter value ϕ is due to sampling fluctuation and any misspecification errors. Thus, the relative biases are the differences in structural estimates after counterbalancing the effects of sampling fluctuation and the existing misspecification errors for the initial model; in other words, they show the effects of adding residual covariances to structural estimates. A negative relative bias value indicates that adding residual covariances tend to decrease the structural estimates comparing to the initial structural estimates. A positive relative bias value, on the other hand, indicates that adding residual covariances increases the structural estimates comparing to the initial structural estimates.

MEASUREMENT MODEL MODIFICATION UNDER MISSPECIFICATION

For each structural parameter, three one-sample t tests were conducted comparing the average relative bias (across the 1,400 replications) to zero; note that these are equivalent to dependent sample t tests comparing biases before ($\hat{\phi}_{before} - \phi$) and after ($\hat{\phi}_{after} - \phi$) modification. Identical analyses were performed for each of the two structural parameters in all 20 cells. The first analysis focused on the relative bias of parameter estimates to detect potential differential bias as a function of parameter value.

Within each condition and for each specific parameter value the variance of the above bias values before and after modification was tested using a test for dependent sample variances (Hinkle, Wiersma & Jurs, 1988). Whereas the first analysis addressed the relative bias of parameter estimates, the second analysis attempted to examine parameter estimate variability. That is, because relative bias estimates can be positive or negative, their average could be near zero, and yet the experimental condition under scrutiny might actually (or additionally) be inducing instability in those estimates. Thus, to evaluate variability in the cells, the following procedures were conducted.

Recall that each cell has 1,400 sets of bias estimates, including those before modification, those after intrafactor residual covariances, those after interfactor residual covariances, and those after both intrafactor and interfactor residual covariances. Within each cell, the 1,400 bias estimates prior to modification were mean centered using their cell mean; this was repeated for the bias estimates intrafactor residual covariances, interfactor residual covariances, and both using their respective cell means. This within-cell centering removed bias and allowed specific focus on variability. Centered values were squared to eliminate sign and these squared values were used in the following analysis:

$$\text{Squared mean centered bias} = (\text{Bias} - \overline{\text{Bias}})^2, \quad (2)$$

where $\text{Bias} = (\hat{\phi} - \phi)$, and $\overline{\text{Bias}}$ is the average of bias for each cell before or after each model modification. Each cell now has 1,400 squared mean centered biases for the initial, intrafactor error covariance, interfactor residual covariance and the combined residual covariances conditions. Three dependent-sample t tests were conducted comparing squared mean centered biases before and after modification. Given the centering and squaring, this is tantamount to dependent-sample tests of variance of the bias estimates. Identical analyses were performed for both structural parameters in all 20 cells.

Eight series of dependent-sample t tests were conducted on the above relative bias estimates to detect the differences of the structural parameter estimate values among the five specified models. First, four dependent-sample t tests were conducted to compare the relative bias estimates in each cell for each of the misspecified model specifications (2, 3, 4 or 5) to those in the corresponding cell of the correctly specified model (specification 1). That is, for every cell for the misspecified models, four dependent-sample t tests were conducted for comparisons between specification 2 versus 1, specification 3 versus 1, specification 4 versus 1, and specification 5 versus 1.

These analyses are to detect potential different effects of adding different residual covariances on the structural parameter estimates for different model misspecification conditions. Similar procedures were also conducted to detect differences on the relative biases of the structural parameter estimate values between specifications 2 and 3, specifications 4 and 5, specifications 2 and 4, and specifications 3 and 5. The purpose is to assess if different measurement model modifications have different impacts with models under different types of misspecification conditions, and also under the same type of misspecification but different degrees of misspecification.

Results

Convergence

Only specification 5 had failures to converge. In 17 of the 20 cells, precisely 1,400 replication attempts were required to obtain

1,400 convergent replications. In the remaining three cells, additional replications were required. For the condition $n = 100$ after modification by adding interfactor residual covariances and the combination of both intrafactor as well as interfactor residual covariances, 1,402 replication attempts were required to yield 1,400. For the condition of $n = 800$ (after modification by adding intrafactor residual covariances), 1,401 replication attempts were required to yield 1,400. Thus, even under the most challenging conditions, the rate of converge was still excellent.

However, in order to ensure that the structural parameter estimates were generated based the same series of data sets across different respecifications and model specification conditions, the three data sets giving nonconvergent replications were eliminated for all the conditions. That is, 1,403 replications were conducted for all conditions yielding 1,400 convergent replications. All analyses were all conducted based on the 1,400 convergent replications.

Description of Modifications

Table 1 shows the average number of modifications of each type made under the different study conditions. Overall, the combination of intrafactor and interfactor residual covariances were added as expected, with interfactor residual covariances only slightly less frequent, followed by intrafactor residual covariances. Note that specification 1 is a correctly specified model and any modifications added are Type I errors due to random sampling fluctuation.

Sample size appeared to have a small but systematic influence for specifications 2 and 3. When the model was misspecified by omitting relevant cross-loading path(s), increases in sample size led to a slight increase in number of the residual covariance respecifications made as the power to identify the misspecification increased. In addition, specifications 2 and 3 with missing relevant paths provided a slightly larger number of residual covariance respecifications than other specifications, increasing by the severity of misspecification. Because paths are omitted to variables that serve as indicators to other factors in specifications 2

and 3, this explains the larger number of modifications for interfactor residual covariances, as well as other residual covariances. Conversely, specifications 4 and 5 with additional irrelevant paths had generally slightly smaller numbers of number of residual covariance respecifications than other specifications. Considering possible other paths explaining the covariances within the model, it is expected that less modifications are to be made.

Average Bias

Table 1 shows the results of average structural parameter estimates before and after any modifications for each condition. For specification 1, the true model is fit to the data. Thus the bias between the estimated structural parameters and true structural parameter is due to sampling error. All modifications suggested by the ML test under this situation are also due to sampling fluctuation, which makes the structural parameter estimates less accurate by slightly increasing the estimates. As sample size increased, the effects of modifications of adding residual covariances tended to decrease as the power increased. It was also noticed that all effects were minimal.

The greatest biases were observed both before and after any modifications for specifications 2 and 3 with missing relevant path(s) and especially for the under-estimated structural parameters. For specification 2, when the path from F1 to V4 is incorrectly fixed to zero, the first structural parameter estimates ($\hat{\phi}_1$) were much more biased than the appropriately estimated structural parameter ($\hat{\phi}_2$). The structural parameter estimates of $\hat{\phi}_1$ were inflated up to 0.07, while the structural parameter estimates of $\hat{\phi}_2$ were inflated to no more than 0.015.

The modifications of residual covariances seemed to help control the inflation for the structural parameter estimates of $\hat{\phi}_1$ to some degree, but with limited effects. For specification 3, when both paths from F1 to V4 and from F2 to V7 are incorrectly fixed to zero, both structural parameter estimates were biased

MEASUREMENT MODEL MODIFICATION UNDER MISSPECIFICATION

Table 1: Average Number of Modifications of Each Type per Replication

Specification	Covariance Condition*	Sample Size			
		100	200	400	800
1	Intra	0.369	0.391	0.404	0.360
	Inter	0.925	0.964	0.968	0.916
	Combine	1.161	1.201	1.208	1.144
2	Intra	1.455	0.564	0.754	1.046
	Inter	1.603	1.631	1.656	1.799
	Combine	1.309	1.513	1.765	1.921
3	Intra	0.650	0.906	1.282	1.701
	Inter	1.141	1.335	1.556	2.190
	Combine	1.559	1.910	2.382	3.114
4	Intra	0.340	0.352	0.363	0.347
	Inter	0.911	0.964	0.968	0.915
	Combine	1.099	1.154	1.159	1.115
5	Intra	0.321	0.336	0.339	0.333
	Inter	0.884	0.936	0.946	0.892
	Combine	1.099	1.154	1.159	1.115

*Note. Intra refers to the intrafactor residual covariance condition, inter refers to the interfactor residual covariance condition, and comb refers to the combination of intrafactor and interfactor residual covariance condition.

and inflated to the similar degree and pattern. For specifications 4 and 5 with irrelevant path(s) - when truly zero path(s) are freely estimated - the structural parameter estimates were moderately inflated at a sample size of 100; as sample size increased, the inflation of the structural parameter estimates became very small.

Relative Bias

Further concerning the impact of different residual covariance modifications on relative bias of estimates for each of the two parameters, two one-sample *t* tests were conducted for each of the 20 cells respectively comparing the average relative bias to zero. Given that two structural parameters appear in each cell, it may be considered that there are two *t* test pairs in each cell, with each of the two parameters being within a model for each modification type. In Table 2, results are presented for the average relative bias

(magnified by 1,000) for each parameter across sample sizes, model specifications and different modification types.

Table 2 shows all relative biases to be no larger than hundredths. Most cells with statistically significant values occurred at specifications 2 and 3, when the model is misspecified with missing relevant paths. Very few significant cells occurred at specifications 1, 4 and 5. For the first structural parameter ϕ_1 at specification 2 when the path from F1 to V4 is missing, all of the relative biases were negative, indicating a propensity for the modification to cause the parameter estimate to become slightly smaller than it is prior to modification.

For the second structural parameter ϕ_2 at specification 2, the significant relative biases were positive when sample sizes are moderate or large, indicating a propensity for the modification to cause the parameter estimate to become slightly larger than it was prior to

modification. The results showed that the structural parameter estimates became smaller and closer to the true parameter value when modifications of residual covariances were made when the initial first structural path is under-represented in the unmodified misspecified model. Conversely, for the initial appropriately estimated second structural path, the effects of modifications were no longer helpful and tended to increase the estimated structural parameter and cause greater biases. However, the practical effects for the second structural parameter were much smaller than for the first under-estimated structural parameter. In addition, it is notable that most of the significant cases occurred when sample sizes were moderate or large.

More significant cells were observed as the severity of misspecification increased when one more relevant path from F2 to V7 is missing at specification 3. Relative biases associated with intrafactor residual covariance modification or the combination of intrafactor and interfactor were almost all negative across sample sizes, whereas in the interfactor case the propensity was for mostly positive relative bias. That is, intrafactor residual covariance or the combination of intrafactor and interfactor modification had a propensity to cause a parameter estimate to become smaller than it was prior to modification.

The interfactor covariance modification, by contrast, had the tendency to cause a

Table 2: Average Structural Parameter Estimates across Modifications of Each Type

Specification	Covariance Condition*	Sample Size							
		100		200		400		800	
		$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$
1	BF	0.610	0.612	0.601	0.605	0.603	0.603	0.602	0.600
	Intra	0.612	0.613	0.603	0.605	0.604	0.603	0.603	0.600
	Inter	0.610	0.613	0.601	0.607	0.604	0.602	0.602	0.600
	Combine	0.612	0.614	0.602	0.607	0.604	0.602	0.603	0.600
2	BF	0.669	0.613	0.660	0.607	0.662	0.604	0.661	0.602
	Intra	0.661	0.607	0.655	0.608	0.653	0.607	0.649	0.608
	Inter	0.662	0.607	0.657	0.607	0.658	0.607	0.658	0.604
	Combine	0.667	0.614	0.659	0.609	0.656	0.608	0.655	0.606
3	BF	0.669	0.665	0.660	0.657	0.662	0.655	0.661	0.653
	Intra	0.668	0.647	0.656	0.627	0.655	0.615	0.652	0.608
	Inter	0.669	0.667	0.662	0.661	0.662	0.659	0.660	0.659
	Combine	0.668	0.650	0.658	0.631	0.656	0.619	0.653	0.611
4	BF	0.614	0.612	0.604	0.605	0.604	0.603	0.602	0.600
	Intra	0.615	0.613	0.606	0.605	0.602	0.603	0.602	0.600
	Inter	0.615	0.612	0.605	0.607	0.604	0.602	0.602	0.600
	Combine	0.617	0.613	0.607	0.607	0.603	0.603	0.602	0.600
5	BF	0.614	0.616	0.604	0.606	0.604	0.603	0.602	0.600
	Intra	0.616	0.615	0.605	0.605	0.603	0.603	0.601	0.600
	Inter	0.618	0.612	0.604	0.607	0.603	0.602	0.602	0.600
	Combine	0.618	0.616	0.605	0.608	0.602	0.603	0.601	0.600

*Note. BF refers to the initial model without any modification, Intra refers to the intrafactor residual covariance condition, inter refers to the interfactor residual covariance condition, and comb refers to the combination of intrafactor and interfactor residual covariance condition.

MEASUREMENT MODEL MODIFICATION UNDER MISSPECIFICATION

parameter estimate to become slightly greater than it was prior to modification. The results indicate that modification to intrafactor residual covariance or the combination of intrafactor and interfactor modification can help make more accurate structural parameter estimates when they are under-represented in the initial model. However, modifications of interfactor residual covariance tend to bias the structural parameter estimates even more under the same situation, although the effects are smaller.

Most of the cells for specifications 2 and 3 with moderate and large sample sizes were significant compared to specification 1, while there were fewer significant cells for specifications 4 and 5 compared to 1. In addition, while comparing the relative biases among specifications 2 to 4 and 3 to 5, many significant results were observed with moderate and large sample sizes. Results showed that misspecifications with missing path(s) had greater impact on the structural parameter estimates after model modifications than misspecifications with additional irrelevant path(s), and comparisons between specifications 2 and 3, and 4 and 5 showed the effects of the severity of the misspecification on the structural parameter estimates after model modifications.

The results have two implications. First, for misspecifications with additional irrelevant paths (specifications 4 and 5), adding one more irrelevant path does not impact the structural parameter estimates to a great degree; second, for misspecifications missing relevant path(s) (specifications 2 and 3), the biases on the structural parameter estimates after the intrafactor or the combined factor model modifications increase as the severity of the misspecification increases; the effects are somewhat ambiguous for interfactor model modification.

Relative Variability

Within each cell, the 1,400 bias estimates prior to modification were mean centered using their cell mean; this was repeated for the bias estimates from the intrafactor residual covariances, the interfactor residual covariances, and the combination of the two. All centered values were squared to eliminate sign;

thus, the average value for a cell was the empirical parameter estimate variance.

In Table 3, the statistical significance of the difference between the original empirical parameter estimate variance and that after each modification is reported for each parameter estimate pair for intrafactor residual covariances, interfactor residual covariances, and the combination of both respectively. Also shown is the average percentage change for each pair of parameter estimates in empirical standard errors after modification, computed as the square root of the empirical parameter estimate variance after modification minus the square root of the empirical parameter estimate variance before modification, divided by the square root of the empirical parameter estimate variance before modification and then multiplied by 100 to create a percentage. That is,

$$\% \text{ change} = \frac{\sqrt{S_{\hat{\phi}_{\text{after}}}^2} - \sqrt{S_{\hat{\phi}_{\text{before}}}^2}}{\sqrt{S_{\hat{\phi}_{\text{before}}}^2}} = \frac{S_{\hat{\phi}_{\text{after}}} - S_{\hat{\phi}_{\text{before}}}}{S_{\hat{\phi}_{\text{before}}}}, \quad (3)$$

where $s_{\hat{\phi}_{\text{after}}}^2$ refers to the empirical parameter estimate variance after modification, $s_{\hat{\phi}_{\text{before}}}^2$ refers to the empirical parameter estimate variance before modification, $s_{\hat{\phi}_{\text{after}}}$ refers to the empirical parameter estimate standard deviation after modification, and $s_{\hat{\phi}_{\text{before}}}$ refers to the empirical parameter estimate standard deviation before modification.

In general, model modification by adding residual covariances appeared to lead to a significant increase in variability of the parameter estimates relative to those prior to modification because more specification errors are introduced into the model. A couple of exceptions occurred in specification 2 with small sample sizes after modification of intrafactor residual covariance or interfactor error covariance, where modification significantly decreased variability of the parameter estimates relative to those prior to modification. Second, empirical standard errors were typically

inflated by less than 8% for specification 1 and varied for other specification conditions when model was misspecified. The most extreme values appeared under specification 2, with inflation reaching between -32.048% and 45.446%. For the other misspecified conditions, the empirical standard error values inflated to no larger than 18%. The modification of interfactor residual covariances generally yielded the smallest inflation in empirical standard errors for specifications 3, 4 and 5.

Conclusion

This study investigated the effects of measurement model modifications of adding residual covariances on structural parameter estimates with different sample sizes under different model specifications. The Monte Carlo

study performed a systematic examination of the impact on the structural parameter estimates for three different common measurement model modifications of residual covariances under five different model specifications, which include adding adding intrafactor residual covariances, adding interfactor residual covariances and adding both intrafactor and interfactor residual covariances respectively. Overall, the model specifications with missing relevant path(s) have the most impact on the structural parameter estimates, while the impact increases as the severity of the misspecification increases.

The propensity is noted that the modifications of either adding intrafactor residual covariance or adding both intrafactor and interfactor residual covariances tended to decrease the structural parameter estimates

Table 4: Percentage Change in Empirical Standard Error of Parameter Estimates across Modifications

Specification	Covariance Condition*	Sample Size							
		100		200		400		800	
		$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_1$	$\hat{\phi}_2$
1	Intra	3.271 ^a	0.527	4.917 ^a	0.899 ^a	5.467 ^a	1.387 ^a	3.769 ^a	1.006 ^a
	Inter	2.191 ^a	2.569 ^a	1.846 ^a	7.222 ^a	2.542 ^a	5.605 ^a	0.992 ^a	3.536 ^a
	Combine	5.345 ^a	2.606 ^a	6.763 ^a	7.608 ^a	7.204 ^a	5.672 ^a	4.356 ^a	3.981 ^a
2	Intra	-29.844 ^a	-31.260 ^a	7.735 ^a	3.942 ^a	10.685 ^a	7.958 ^a	8.230 ^a	6.107 ^a
	Inter	-31.114 ^a	-32.048 ^a	-15.347 ^a	-15.207 ^a	7.250 ^a	6.187 ^a	1.932	1.592
	Combine	5.441 ^a	3.342 ^a	9.161 ^a	6.896 ^a	9.802 ^a	7.281 ^a	45.446 ^a	40.501 ^a
3	Intra	4.213 ^a	5.781 ^a	8.109 ^a	11.294 ^a	9.719 ^a	16.024 ^a	7.437 ^a	14.323 ^a
	Inter	2.228 ^a	0.941	2.715 ^a	2.709 ^a	2.777 ^a	2.185 ^a	4.145 ^a	2.987 ^a
	Combine	5.423 ^a	6.326 ^a	9.037 ^a	12.760 ^a	9.422 ^a	17.242 ^a	8.772 ^a	16.700 ^a
4	Intra	4.596 ^a	0.341	10.827 ^a	0.914 ^a	15.636 ^a	1.339 ^a	15.115 ^a	1.126 ^a
	Inter	2.174 ^a	2.703 ^a	2.359 ^a	7.280 ^a	1.995 ^a	5.915 ^a	1.473 ^a	3.633 ^a
	Combine	6.239 ^a	2.610 ^a	10.134 ^a	7.721 ^a	13.483 ^a	6.367 ^a	12.822 ^a	4.047 ^a
5	Intra	3.969 ^a	-1.289	11.371 ^a	0.357	15.096 ^a	0.666 ^a	17.261 ^a	0.435
	Inter	4.430 ^a	-5.187 ^a	4.168 ^a	3.043 ^a	5.925 ^a	2.303 ^a	3.274 ^a	1.140
	Combine	7.939 ^a	-2.998 ^a	12.647 ^a	4.016 ^a	17.311 ^a	3.892 ^a	14.187 ^a	2.314 ^a

*Note. Intra refers to the intrafactor residual covariance condition, inter refers to the interfactor residual covariance condition, and comb refers to the combination of intrafactor and interfactor residual covariance condition.

^aThe empirical parameter estimate variance was statistically significantly different (p<.05) from the corresponding variance prior to model modification.

MEASUREMENT MODEL MODIFICATION UNDER MISSPECIFICATION

compared to those prior to modification, while the interfactor modification tended to increase the structural parameter estimates. By contrast, the model specifications with additional irrelevant path(s) did not have much impact on the structural parameter estimates because very few random significant cells were observed at specifications 4 and 5. In sum, only the underestimated structural parameter estimates from the misspecified models with missing relevant path(s) approached the true values when the intrafactor residual covariances or the combination of intrafactor and interfactor residual covariances were added. In addition, all three model modifications tended to significantly increase the variability of the parameter estimates relative to those prior to modification across model specification conditions. Interestingly, sample size did not appear to influence the impact of the measurement model modifications of residual covariances on the structural parameter estimates. Thus, more attention should be paid to a misspecified model with missing path(s) when conducting measurement model modification.

Although many significant cells were observed, the impacts of the different types of model modification were found to be small, usually no larger than hundredths. However, it is notable that this study focused on conditions with normally distributed data. It is expected that the impact of measurement model modification would be enlarged under the inevitable conditions with nonnormally distributed data in practical research. Based on results from this investigation it may be stated that liberally saturating measurement models with additional residual covariance parameters does not necessarily help with the structural parameter estimates if a researcher begin with a misspecified model, especially with missing parameters. Thus, the role of theory cannot be more salient for all model types (Hancock, 1999).

However, researchers should exercise caution with the results. As with any simulation study, there are innumerable conditions to manipulate and choices must be made to keep the design manageable. It is possible that another research study could produce different results under different models and experimental

conditions. For example, the extent to which measurement model modification alters structural parameter estimates under conditions of nonnormality remains to be studied. Similarly, the amount of modification that affects parameters within a mean structure, such as latent means and intercepts, is also a subject for further investigation.

References

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411-423.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino, CA: Multivariate Software.
- Chou, C.-P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange multiplier, and Wald tests. *Multivariate Behavioral Research, 25*, 115-136.
- Chou, C.-P., & Bentler, P.M. (2002). Model modification in structural equation modeling by imposing constraints. *Computational Statistics & Data Analysis, 41*, 271-287.
- Fan, W., & Hancock, G. R. (2006). Impact of post hoc measurement model over-specification on structural parameter integrity. *Educational and Psychological Measurement, 66*, 748-764.
- Green, S. B., & Babyak, M. A. (1997). Control of Type I error with multiple tests of constraints in structural equation modeling. *Multivariate Behavioral Research, 32*, 39-52.
- Green, S. B., Thompson, M. S., Babyak, M. A. (1998). A Monte Carlo investigation of methods for controlling Type I errors with specification searches in structural equation modeling. *Multivariate Behavioral Research, 33*, 365-384.
- Green, S. B., Thompson, M.S., & Poirier, J. (1999). Exploratory analyses to improve model fit: Errors due to misspecification and strategy to reduce their occurrence. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 113-126.

- Green, S. B., Thompson, M. S., & Poirier, J. (2001). An adjusted Bonferroni method for elimination of parameters in specification addition searches. *Structural Equation Modeling: A Multidisciplinary Journal*, 8, 18-39.
- Hancock, G. R. (1999). A sequential Scheffé-type respecification procedure for controlling Type I error in exploratory structural equation model modification. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 158-168.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1988). *Applied statistics for the behavioral sciences*. Boston: Houghton Mifflin Company.
- Hutchinson, S. R. (1993). Univariate and multivariate specification search indices in covariance structure modeling. *The Journal of Experimental Education*, 61, 171-181.
- Hutchinson, S. R. (1998). The stability of Post Hoc model modifications in confirmatory factor analysis models. *The Journal of Experimental Education*, 66, 361-380.
- Kaplan, D. (1988). The impact of specification error on the estimation, testing and improvement of structural equation models. *Multivariate Behavioral Research*, 23, 69-86.
- Kaplan, D. (1989). Model modification in covariance structure analysis: Application of the expected parameter change statistic. *Multivariate Behavioral Research*, 24, 285-305.
- Kerlinger, F. (1986). *Foundations of behavioral research (3rd Ed.)*. New York: Freeman.
- Kline, R. B. (2004). *Principles and practice of structural equation modeling (2nd Ed.)*. New York: Guilford.
- MacCallum, R. C. (1986). Specification searches in covariance structure analysis. *Psychological Bulletin*, 100, 107-120.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490-504.
- Mulaik, S. A., & Millsap, R. E. (2000). Doing the four-step right. *Structural Equation Modeling: A Multidisciplinary Journal*, 8, 36-73.
- Newcomb, M. D., & Bentler, P. M. (1988). Impact of adolescent drug use and social support on problems of young adults: A longitudinal study. *Journal of Abnormal Psychology*, 97, 64-75.
- Paxton, P., et al. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling: A Multidisciplinary Journal*, 8, 287-312.
- Schumacker, R. R., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling (2nd Ed.)*. Mahwah, NJ: Lawrence Erlbaum.
- Silvia, E. S., & MacCallum, R. C. (1988). Some factors affecting the success of specification searches in covariance structure modeling. *Multivariate Behavioral Research*, 23, 297-326.
- Thompson, B., Cook, C., & Heath, F. (2003). Structure of perceptions of service quality in libraries: A LibQUAL+™ study. *Structural Equation Modeling: A Multidisciplinary Journal*, 10(3), 456-464.