

# STRATEGIES FOR DEALING WITH MEASUREMENT ERROR IN MULTIPLE REGRESSION

Robert A. McDonald, Rensselaer Polytechnic Institute, Troy, New York, USA  
Scott J. Behson, Fairleigh Dickinson University, Teaneck, New Jersey, USA  
Charles F. Seifert, Siena College, Loudonville, New York, USA

## ABSTRACT

*Methods for handling measurement error in regression analysis range from assuming that measures are perfectly reliable to explicitly modeling measurement error through the use of multiple indicators. In two simulation studies, we examined the effectiveness of several strategies for incorporating measurement error in regression. First, we compared four strategies: (1) multiple indicator structural equation modeling, (2) a composite indicator structural equation (CISE) model with adjustment for measurement error based on Cronbach's alpha of the composite, (3) a composite indicator structural equation model with no adjustment for measurement error, and (4) a single indicator model with no adjustment for measurement error. The second study explored the consequences of either over- or under- estimating measurement error while using the CISE alpha method.*

**Keywords:** *Structural Equation Modeling, Measurement Error, Monte-Carlo Simulation, Multiple Regression*

## 1. INTRODUCTION

Multiple regression has become an important multivariate data analytic technique in the management and industrial organizational fields. Over the past two decades, increased attention has been paid to the issue of measurement error (Anderson, Stone-Romero & Tisak, 1996) and representation of constructs as latent variables (Bagozzi & Edwards, 1998). The recognition that virtually all social science depends on fallible measures is extremely important because biased parameter estimates and standard errors can occur when measurement error is ignored in regression analysis (Hayduk, 1987).

## 2. HANDLING MEASUREMENT ERROR

Strategies designed to address the measurement error issue in regression situations include simultaneous estimation of measurement error and regression parameters, estimation of measurement error and regression parameters in separate analyses, and fixing measurement error to an *a priori* value. Multiple indicator structural equation (MISE) modeling using maximum likelihood estimation is an example of a commonly used strategy for simultaneously estimating measurement error and regression parameters (Jöreskog & Sörbom, 1993). The major advantage of this technique is that all the information in the variance-covariance matrix of the variables under consideration is used in the simultaneous estimation of measurement error and regression parameters. Using all available information can increase the efficiency of parameter estimates. However, specification errors or estimation errors in one part of the model can cause biased parameter estimates throughout the model as well as convergence problems (Bollen, 1996). A second approach is the estimation of measurement error of an indicator independent of the estimation of the regression parameters (Bagozzi & Edwards, 1998, p. 57). Specifically, the estimation of measurement error associated with each indicator is done on an indicator-by-indicator basis, not simultaneously. Mis-specification or mis-estimation of measurement error for one indicator will not affect estimates of measurement error for another indicator. A third approach for dealing with measurement error is to fix the proportion of measurement error for an indicator to an *a priori* value (Hayduk, 1987, p. 120). This approach has the advantage that the estimation of measurement error is based on a different sample than the sample used for the estimation of the regression parameters. Thus, the estimation of the measurement error and the regression parameters are independent. In this study, several methods for handling measurement error are compared.

### 2.1. Single-Indicator Approach

With traditional standard regression procedures, variables are assumed to be perfectly reliable without contamination from measurement error. The problem with this assumption is that measurement error can

attenuate estimated relationships among latent variables and bias standard error terms associated with the estimated relationships. The overall effect of not incorporating measurement error in data analysis can be a reduction in statistical power for detecting relationships among variables. Since this technique has many known drawbacks, many have advocated for alternate methods that incorporate measurement error.

## 2.2. Multiple Indicator Structural Equation Modeling

One commonly used data analytic technique for the simultaneous estimation of measurement error and regression parameters is multiple indicator structural equation modeling (MISE) with maximum likelihood estimation. This model is the *total disaggregation model* in the construct framework proposed by Baggozi and Edwards (1998). MISE is a powerful data analytic technique for estimating simultaneous linear regression equations for latent variables in the context of a measurement error theory (Jöreskog & Sörbom, 1993). With the MISE approach, structural and measurement models are specified. The structural model specifies the relationship among the latent variables and the measurement model specifies the relationship of the latent variables with the indicators and the error structure of the indicators. For example, the structural model specifying a situation with three latent exogenous variables related to a latent endogenous variable is defined as:

$$L_Y = \beta_1 L_{X1} + \beta_2 L_{X2} + \beta_3 L_{X3} + \varepsilon \quad [1]$$

where  $\beta_1$  is the path coefficient reflecting the unique relationship between  $L_{X1}$  and  $L_Y$  in the presence of the other predictors,  $\beta_2$  is the path coefficient reflecting the unique relationship between  $L_{X2}$  and  $L_Y$  in the presence of the other predictors,  $\beta_3$  is the path coefficient reflecting the unique relationship between  $L_{X3}$  and  $L_Y$  in the presence of the other predictors,  $\varepsilon$  is the residual error variance.

The measurement model form relating  $L_Y$  to its indicators is the traditional factor analytic model, which in the case of three indicators, can be expressed as follows:

$$Y_1 = \alpha_1 + \lambda_1 L_Y + \varepsilon_1 \quad [2]$$

$$Y_2 = \alpha_2 + \lambda_2 L_Y + \varepsilon_2 \quad [3]$$

$$Y_3 = \alpha_3 + \lambda_3 L_Y + \varepsilon_3 \quad [4]$$

where  $Y_k$  is an observed measure of  $L_Y$ ,  $\alpha_k$  is an intercept term,  $\lambda_k$  is a regression (or path) coefficient and  $\varepsilon_k$ . Common assumptions in applications of this formulation are that  $L_Y$  and the various  $\varepsilon_k$  are continuous and normally distributed with a mean of zero. One of the  $\lambda_k$  is fixed at 1.0 in order to establish a metric for  $L_Y$ , and the various intercepts are also fixed at an arbitrary value in order to facilitate model identification. The relationship of the other latent variables and their associated indicators are specified in a similar manner. Maximum likelihood methods are then applied to estimate the population covariance matrix for  $L_Y$ ,  $L_{X1}$ ,  $L_{X2}$  and  $L_{X3}$  and this, in turn, is used to estimate  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$ . Estimated standard errors are also derived as well as significance tests (see Joreskog and Sorbom, 1996, for a description of the general underlying theory).

The structural model specifies the relationships among latent variables. Both the measurement and structural models are simultaneously estimated such that structural relations among latent variables are corrected for attenuation due to measurement error. MISE techniques, such as maximum likelihood estimation, simultaneously estimate model parameters using all information in the variance-covariance matrix. Another advantage of the MISE approach is that the structure of measurement model and potential mis-specification in structure of the indicators can be examined (Baggozi & Edwards, 1989; Little et al., 2002). Modification indices and residuals can be inspected to reveal points of stress in the model.

Despite the advantages of the MISE approach, there are several drawbacks to this method. First, the underlying assumptions of MISE are asymptotic (Jöreskog & Sörbom, 1993); that is, this method is effective only with large samples. Problems with convergence and offending estimates are particularly troublesome with small samples and with complex models with a large number of parameters (Tanaka, 1987). With a MISE approach, regression equations are required to specify measurement and structural portions of the model. In contrast, standard multiple regression only has a structural model, not an explicit measurement model. Consider a regression situation with one dependent variable and three predictors, each with three indicators. There are 30 estimated parameters for a multiple indicator model as compared to only 10 with standard multiple regression. A second disadvantage is that specification

errors in one part of the model can bias parameter estimates throughout the model (Bollen, 1996). Any mis-specification in the equations relating a latent variable with associated indicators can affect the estimation of measurement parameters linking the remaining latent variables to associated indicators.

### 2.3. Composite Indicator Structural Equation Modeling

To overcome some of the inherent problems with the MISE method, some researchers have used single composite indicators structural equation modeling approach (CISE). In this approach, measurement error ( $\sigma_{\epsilon 1}^2$ ) for the composite indicator is fixed to an estimate of the measurement error (Hayduk, 1987). Typically the fixed value for measurement error is either based on an estimate of reliability (such as from the formula for Cronbach's alpha or split-half reliability) or on *a priori* knowledge of the reliability of the measure. This approach for handling measurement error is analogous to errors-in-variables regression (Heise, 1975, 1986; Fuller & Hidiroglou, 1978; Warren, Keller-White, & Fuller, 1974) in which regression coefficients are estimated in the presence of errors in measurement. Specifically, an estimate of measurement error for each variable is used to correct the variance-covariance matrix for attenuation.

Both psychometric advantages and model advantages of bundling items into parcels (such as a single composite indicator) have been noted by researchers (Little et al., 2002). Psychometric advantages tend to focus on the tendency of a composite scale to be more reliable, more normally distributed, and have more intervals between scale points than a single indicator (Bandalos, 2002). Rushton, Brainerd and Pressley (1983) argued that individual items are less reliable than composite scores and demonstrated with published data that relations between theoretically linked constructs were detected with composite scores but not with individual items. The improvement in reliability occurs because measurement errors in an item are offset by measurement errors in other items. For example, three items representing a latent construct each with an item reliability of .7 will yield a Cronbach's alpha of .88 when combined into a single composite scale.

Little et al. (2002, p. 157) noted that "item distributions, which may have problems with skewness and kurtosis (thus violating assumptions of statistical inference), become more normally distributed when aggregated into scale scores or parcels." By aggregating items, deviations from a normal distribution in one item may be compensated for by deviations from a normal of another item. For example, one item may have a right skew and another item may have a left skew. By combining these two items into a composite scale, the two skews offset each other to yield a composite that is normally distributed. Non-normality has been found to adversely affect standard errors and fit indices in structural equation modeling (Finch, West, & MacKinnon, 1997; Muthén & Kaplan, 1985).

Several researchers have recognized that more intervals between scale points can occur when aggregating data from items into a composite scale which data that closer approximate continuous properties than the separate items that comprise the composite scale (Little et al., 2002). For example, aggregating two items each measured with 3-point response scales yields a composite indicator with 5 scale points. This is important because scale coarseness has been found to bias regression (Russell & Bobko, 1992; Russell, Pinto, & Bobko, 1991) and structural equation modeling results (Babakus, Ferguson, & Joréskog, 1987; DiStefano, 2002; Green et al, 1997; Huchinson & Olmos, 1998).

Cited model advantages of composite scales include Type I errors, isolation of specification errors, solution stability, and identification issues. The use of composite scales reduces the probability of detecting a spurious relationship by chance, because a smaller variance-covariance matrix is used as input information for estimation. As noted by Little et al. (2002, p. 160), "a model with three constructs each measured with 10 variables would yield about 22 spurious correlations" using a .05 Type I error rate. However, a structural equations model with three constructs each measured with one variable, a 3-variable correlation matrix with  $3(2)/2 = 3$  unique correlations, would contain less than one spuriously significant correlation.

Items in a composite scale likely share specific sources of variance which reflect systematic source of common variance that are likely will not be modeled in multiple indicator structural equation model (Little et al, 2002). Examples of shared variance include common method variance and social desirability bias. Little et al. (2002, p. 161) claimed that aggregating items "would likely eliminate or at least reduce the

unwanted source or sources and would lead to a better model fit than if the items were used as indicators of the construct.”

Solution stability is an advantage commonly proposed for the use of aggregated items in a structural equation model. Fewer parameters are estimated when a single composite scale is used as an indicator of a latent variable than when multiple indicators are used. Only structural parameters are estimated because measurement error variances for the indicators of the latent variables are fixed. Fewer estimated parameters may be particularly beneficial with small sample sizes because structural equation modeling is based on asymptotic assumptions. Having fewer estimated parameters reduces sample size requirements (Hall, Snell, & Foust, 1999, Tanaka, 1987).

The final model advantage of using a composite scale with fixed measurement error rather than individual indicators is related to identification of the model (Little et al., 2002). Using a composite scale with fixed measurement error leads to a just-identified model that has only one unique solution. A MISE model is an overidentified model that can have more than one optimal solution.

The disadvantage of the CISE method is that accurate estimates of the fixed measurement error variances must be specified. One option is to estimate and fix measurement error associated with each composite scale. This can be accomplished by multiplying the proportion of measurement error variance by the variance of the composite scale. The proportion of the measurement error variance is the difference between 1 and an estimate of the reliability of the composite scale calculated using a formula for measurement reliability, such as Cronbach's alpha or test-retest coefficient. A disadvantage of this option is that the estimate of measurement error is based on the same sample that is used to test the structural model. Methods for non-simultaneous estimation of measurement error and structural parameters, such as errors-in-variables approach, were designed specifically for situations where the estimates of measurement error are obtained from a source independent of the sample (Fuller & Hidiroglou, 1978). A second option is to use *a priori* estimates of the reliability of the composite scales. Although an independent estimate of measurement error is obtained, *a priori* estimates of the reliability of a composite scale are only valid for samples that closely approximate the original sample from which the *a priori* estimates were estimated. Biased parameter estimates will be yielded when inappropriate *a priori* estimates of measurement error are used.

As with the MISE approach, model misspecification can degrade parameter estimates and standard errors. However, unlike the MISE approach, the structure of the individual indicators cannot be examined. Thus, misspecification in the underlying measurement model may be hidden with the CISE approach (Little et al., 2002). With the MISE approach, modification indices and residuals within the measurement portion of the model can be examined for evidence of model misspecification. Finally, the CISE alpha model reduces the possibility of misfit affecting the entire measurement model. Thus, it is possible that the use of this method could result in spuriously high levels of model fit, as compared to models that incorporate potential misfit in both the structural and measurement models.

### 3. STUDY 1

Monte Carlo simulation methodology was used to examine the effectiveness of four strategies for incorporating measurement error in regression in a case with three latent exogenous variables related to a latent endogenous variable as specified in Equation [1]. The  $\beta_1$  in Equation [1] was manipulated across specific levels and used as the focal path for this study. Although  $\beta_2$  and  $\beta_3$  were manipulated, the manipulation of these parameters was chosen to obtain specific values for  $\beta_1$  and  $\varepsilon$  in Equation [1].

Four data analytic techniques were examined in Study 1. The first data analytic technique was a full structural equation model where both the structural and measurement parameters were simultaneously estimated in a multiple indicator structural equation model (MISE). The second data analytic technique (CISE alpha) was a composite single indicator model with the measurement error parameters fixed based on reliability estimates gained from the formula for Cronbach's alpha. The third data analytic technique was a composite single indicator model with the no measurement error (CISE no error) and was included to reflect a regression approach in which composite scales assumed to be infallible are used in the

regression model. The fourth data analytic technique was a single indicator model with measurement error parameters fixed at zero (SIM). We included this latter data analytic technique because this reflects the commonly used regression approach where the researcher assumes that a single infallible measure can adequately assess the scores of individuals on a construct.

### 3.1. Design Considerations

PRELIS was used to generate data with known statistical properties. Data were generated using 2,000 replications. The following data characteristics were manipulated: sample size, effect size of  $\beta_1$ , residual error variance, number of indicators, and reliability of the composite scale. We conducted a review of all studies that used multiple regression or structural equation modeling in the *Journal of Management* from 1994 to 1998 to ensure the representativeness of the design parameters. The sample size was manipulated with three levels: 75, 150, and 300. These sample sizes closely correspond to the 25<sup>th</sup>, 50<sup>th</sup>, and 75<sup>th</sup> percentiles for the sample sizes we found in our review. The effect size for  $\beta_1$  was conceptualized in terms of partial *eta*-squared in which the parameter reflected an effect size of either 0% additional explained variance over and above the other predictors (to evaluate Type I errors), 5% additional explained variance, or 10% additional explained variance. The five percent partial *eta*-squared reflects a medium effect and the 10 percent partial *eta*-squared reflects a strong effect. The residual error variance was manipulated with three levels: 30, 50, and 70 percent. These levels of residual error variance were representative of those found in our review of articles in the *Journal of Management*. The number of indicators was manipulated with two levels: three and six indicators. In a review of scales appearing in 75 articles published in several management-oriented journals from 1989 through 1993, Hinkin (1995) found that approximately sixty percent of the scales had three to six items. The reliability of the composite scale was manipulated with two levels: .7 and .9. In his review, Hinkin (1995) found that internal consistency reliabilities ranged from .55 upwards with only 12 percent of the scales with reliabilities lower than .7.

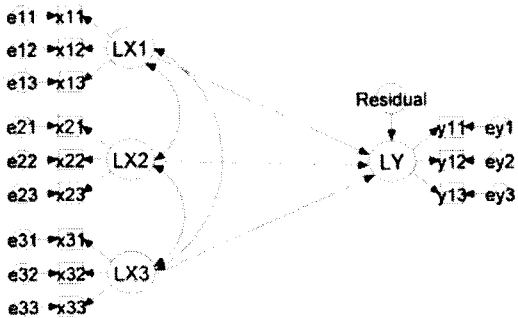
### 3.2. Generation of Data

The simulation involved generating data for the latent variables that conformed to the structural portion of the study design and then generating data for the measurement portion. First, three predictors inter-correlated at the .2 level were generated. These three predictors had a population mean of zero and unit variance. A latent dependent variable with a population mean of zero and unit variance was generated using equation [1]. The effect size for  $\beta_1$  was conceptualized in terms of partial *eta*-squared in which the parameter reflected an effect size of either 0% additional explained variance over and above the other predictors (to evaluate Type I errors), 5% additional explained variance, or 10% additional explained variance. The other regression parameters were set to values so that the desired partial *eta*-squared for  $\beta_1$  and residual error variance were obtained. The indicators for the latent variables were generated by use of Equations [2] to [4]. Values for  $\lambda_k$  were determined such that the composite scale had the desired Cronbach's alpha and a population mean of zero and unit variance.

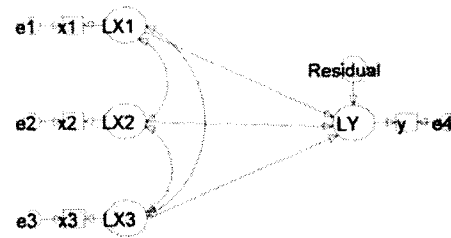
### 3.3. Method of Analyses

The data conditions were analyzed by four data analytic techniques, 1) multiple indicator structural equation model (MISE), 2) a composite indicator structural equation model with adjustment for measurement error (CISE alpha), 3) a composite indicator structural equation model with no adjustment for measurement error (CISE no error), and 4) a single indicator model with no adjustment for measurement error (Please see Figures 1 through 4 to see the models we developed for each technique). In the multiple indicator structural equation model, the measurement and structural parameters were simultaneously estimated. The indicators each had a unique path from the appropriate latent variable and the measurement error for each indicator was freely estimated. For the second method, a composite single indicator model with the measurement error parameters fixed based on reliability estimates gained from the formula for Cronbach's alpha was estimated. The value for the fixed measurement parameter was determined by first calculating the Cronbach's alpha for the composite scale. Then this reliability estimate was subtracted from one and then multiplied by the variance of the composite scale. For the third method, composite single indicator model with measurement error parameters fixed to zero was estimated. Finally, a single indicator model with measurement error parameters fixed at zero was estimated.

**FIGURE 1. MISE MODEL**

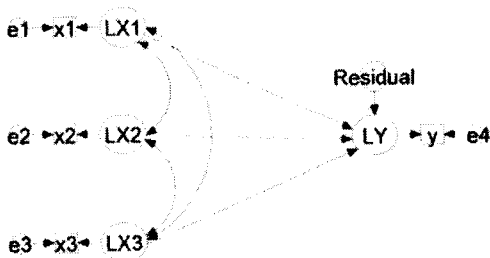


**FIGURE 2. CISE ALPHA MODEL**

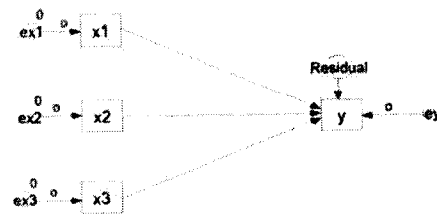


Note: The variances of e1, e2, and e3 were set to equal one minus the reliability of the composite scale times the variance of the composite scale ((1-alpha)\*variance).

**FIGURE 3. CISE NO ERROR MODEL**



**FIGURE 4. SINGLE INDICATOR MODEL**



Note: The measurement error represented by e1, e2 and e3 was assumed to be zero.

#### 4. RESULTS

The focus of the results is on the estimation of the  $\beta_1$  path coefficient. The results concerning the estimation of the path coefficient are based only on valid solutions in a data condition. Consistent with other Monte Carlo simulation studies (e.g., Xitao, Thompson, & Wang, 1999), solutions that did not converge and solutions containing offending estimates were eliminated from data analysis. The maximum number of iterations was set to 1,000 for each solution.

Average bias in the estimation of the  $\beta_1$  path coefficient across replications of a Type II data condition was calculated using the following formula<sup>1</sup>:

$$APB = (\sum (\beta_1 / B_1) / n) * 100 \quad [5]$$

where APB is the average bias in the parameter estimates of  $\beta_1$  across  $n$  replications for a Type II data condition,  $B_1$  is the true population value, and  $\beta_1$  is the estimated parameter. Average bias values greater than 100 percent reflect over-estimation of the parameter. Because the true population value for  $\beta_1$  was zero in the Type I error conditions, the average error between the true population value and the estimated  $\beta_1$  parameter across replications was used as an index of bias in estimated  $\beta_1$  parameters. Average parameter error was calculated using the following formula:

$$APE = \sum (\beta_1 - B_1) / n \quad [6]$$

where APE is the average error in estimating the  $\beta_1$  parameter across  $n$  replications. Negative average error values represent underestimation of the parameter.

To provide a perspective on the accuracy of the standard errors, we first calculated the root mean square error (RMSE) of the parameter estimates. RMSE is an empirical estimate of the mean standard error across  $n$  replications and was calculated using the following formula:

$$RMSE = (\sum (\beta_1 - B_1)^2 / n)^{.5} \quad [7]$$

where RMSE is the root mean square error across  $n$  replications. Then we calculated an index of average standard error bias for the  $\beta_1$  parameter estimate by use of the following formula:

$$\text{ASEB} = (\text{se} / \text{RMSE}) * 100 \quad [8]$$

where ASEB is the average bias in the standard error for the  $\beta_1$  parameter estimate across the  $n$  replications and  $\text{se}$  is the average standard error across the  $n$  replications. An average standard error bias value greater than 100 percent represents an inflated standard error.

To evaluate Type I errors, we calculated the proportion of cases where the null hypothesis of a zero  $\beta_1$  was incorrectly rejected. To evaluate Type II errors, we calculated the percentage of cases where the null hypothesis of a non-zero  $\beta_1$  was correctly rejected (i.e., we calculated the power of the test). Because the sample sizes are small relative to the asymptotic assumptions of maximum likelihood, we decided to use the  $t$  distribution to define critical values in all cases (recognizing that, in fact, the nature of the sampling distributions are intractable in maximum likelihood analysis with small sample sizes). We also calculated the proportion of times that the true population parameter value occurred within the 95 percent confidence interval.

#### 4.1. Type I Error Conditions

*Offending estimates and non-converging solutions.* Averaged across all the data conditions, the MISE model had the most solutions with offending estimates and non-converging solutions. On average per 2,000 replications, the MISE model had 62.3 solutions with offending estimates and 0.8 solutions that did not converge within 1,000 iterations. Averaged across all the data conditions, the CISE model with alpha had 12.9 solutions per 2,000 replications with offending estimates and no solutions that did not converge. The CISE no error model and the single indicator model had no solutions with offending estimates because these models underestimated parameters. Further, all solutions converged for the CISE no error model and the single indicator model because these models were just-identified.

For both the full model and the CISE alpha model, the average number of solutions with offending estimates was higher for all the data conditions with .6 reliability than for the data conditions with .8 reliability. For the MISE model, the average number of solutions with offending estimates for the .6 reliability data condition was 112.8 and for the .8 reliability condition was 11.72. The average number of solutions with offending estimates for the .6 reliability data condition was 25.72 and for the .8 reliability condition was 0.01 for the CISE alpha model. The number of solutions with offending estimates increased with decreasing sample size for the full model and the CISE alpha model. For the MISE model, the average number of solutions with offending estimates decreased from 130.8 for the 75 sample size data condition to 14.58 for the 300 sample size data condition.

Non-converging solutions with the MISE model occurred with the smallest sample size and the lowest reliability data conditions. Specifically, the average number of non-converging solutions per 2,000 replications drop from 2.5 for the .7 reliability data conditions to zero for the .8 reliability data conditions. The average number of non-converging solutions per 2,000 replications was 2.3 with a sample size of 75 and zero for the larger sample size data conditions. For the two three and six item data conditions and the three residual error conditions, the average number of non-converging solutions per 2,000 replications ranged from 0.6 to 0.8.

*Parameter estimate error and standard error bias.* Table 1 presents the average parameter estimate error and standard error bias for each level of the design factors and averaged across all the data conditions for the Type I error conditions. Averaged across all the data conditions, both the MISE model (APE = -.002) and the CISE alpha model (APE = -.002) had parameter estimates of the  $\beta_1$  path coefficient that were more accurate than the CISE no error model (APE = .025) and the single indicator model (APE = .037). Average error in the estimate of the  $\beta_1$  parameter increased with decreasing reliability and increasing number of indicators for all the models. For the MISE model and the CISE alpha model, the data conditions with a sample size of 75 had the highest amount of error in the  $\beta_1$  parameter estimate. For all the models, the highest average error in the estimate of the  $\beta_1$  parameter across the residual error conditions was for the 30 percent residual error condition.

**TABLE 1: AVERAGE PARAMETER ESTIMATION ERROR, STANDARD ERROR BIAS AND TYPE I ERROR RATE FOR TYPE I ERROR CONDITIONS**

Data Conditions	MISE Model			CISE Alpha Model			CI No Error Model			Single Indicator Model		
	APE	ASEB	Type I Error	APE	ASEB	Type I Error	APE	ASEB	Type I Error	APE	ASEB	Type I Error
<u>Reliability</u>												
.70	-.003	92.4	4.1	-.004	98.1	4.9	.034	86.0	8.7	.038	88.6	8.4
.90	-.001	97.6	5.0	-.001	98.3	5.2	.015	96.3	5.9	.035	87.6	8.8
<u>Items</u>												
3	-.001	95.0	4.8	-.002	98.5	5.1	.024	90.4	7.1	.036	87.9	8.6
6	-.002	95.0	4.3	-.003	97.9	5.1	.025	91.9	7.5	.037	88.2	8.5
<u>Residual Error</u>												
30%	-.002	96.1	4.3	-.004	98.2	4.9	.029	88.9	8.5	.044	84.2	10.2
50%	-.002	94.4	4.6	-.001	97.9	5.3	.025	89.7	7.3	.037	87.5	8.6
70%	-.002	94.5	4.7	-.002	98.5	5.0	.020	94.9	6.2	.029	92.5	6.9
<u>Sample Size</u>												
75	-.003	91.0	4.3	-.005	97.7	4.6	.025	92.9	6.1	.037	92.6	6.8
150	-.001	95.8	4.4	-.001	98.0	5.3	.025	92.6	7.0	.036	89.1	8.1
300	-.001	98.1	4.9	-.002	98.9	5.3	.025	88.3	8.9	.036	82.4	10.8
Overall Mean	-.002	95.0	4.5	-.002	98.2	5.1	.025	91.2	7.3	.036	88.1	8.6
Baseline	0	95	5	0	95	5	0	95	5	0	95	5

Averaged across all the data conditions, the CISE alpha model had the least bias in the standard errors (ASEB = 98.2) associated with the  $\beta_1$  parameter estimate, followed by the MISE model (ASEB= 95.0), the CISE no error model with (ASEB = 92.3), and the single indicator model (ASEB = 88.1). Average standard error bias decreased with increasing sample size for the full model and CISE alpha model. For the CISE model with no error and single indicator model, average standard error bias decreased with decreasing sample size. Standard errors were more accurate for the .9 reliability conditions than for the .7 reliability conditions for all the models, except for the single indicator model.

*Type I error rate.* Table 1 also presents the average Type I error rate for each level of the design factors and averaged across all the data conditions for the Type I error conditions. Averaged across all the data conditions, the Type I error rate for the CISE alpha model (Type I error rate = 5.1%) most closely approximated the nominal rate of 5%, followed by the MISE model (Type I error rate = 4.5%). Both the CISE no error model (Type I error rate = 7.3%) and the single indicator model (Type I error rate = 8.6%) had liberal Type I error rates.

The average Type I error rate for the MISE model most closely approximated the nominal rate of 5.0% with higher item reliability, fewer items, larger residual error, and larger sample size. For the CISE alpha model, the average Type I error rate reasonably approximated the nominal rate across all data conditions. Liberal average Type I error rates were obtained across all data conditions for the CISE no error model and the single indicator model with no average Type I error rate falling below 5.9% for either model in any of the data conditions.



## 4.2. Type II Error Condition

*Offending estimates and non-converging solutions.* The pattern of the results for the Type II data conditions are similar to the Type I error conditions. Specifically, the MISE model had 58.2 solutions per 2,000 replications with offending estimates and the CISE alpha model had 9.8 offending solutions per 2,000 replications. The MISE model was the only method with non-converging solutions, 0.1 solutions per 2,000 replications. The CISE no error model and the single indicator model had no solutions with offending estimates. The pattern for the average number of solutions with offending and non-converging solutions across the data conditions are similar to the Type I error conditions.

*Parameter estimate and standard error bias.* Table 2 presents the average parameter estimate error and standard error bias for each level of the design factors and averaged across all the data conditions for the Type II error conditions. Averaged across all data conditions, the average parameter estimate bias was trivial for the CISE alpha model (APB = 100.5) and for the MISE model (APB = 101.8). Across the data conditions, the single indicator model (APB = 61.9) and the CISE no error (APB = 86.6) tended to severely underestimate  $\beta_1$ . There were only trivial differences in average parameter estimate bias for the full model and the CISE alpha model across the data conditions. For the CISE no error model with and the single indicator model, higher reliability and smaller residual error was associated with less average parameter estimate bias. Conditions with more items and larger effect sizes were associated with more parameter bias.

The least amount of standard error bias was obtained with the CISE alpha model (ASEB = 98.1), followed by the MISE model (ASEB = 94.4), the CISE no error model (ASEB = 85.2), and the single indicator model (ASEB = 62.9). There were only trivial differences in average standard error bias for the CISE alpha model across the data conditions. For the MISE model, two noteworthy differences across the data conditions were apparent. First, the average standard error bias dropped from 97.7 for the .9 reliability data condition to 93.1 for the .7 reliability data condition. Second, average standard error bias ranged from 91.0 for the sample size of 75 data conditions to 98.4 for the sample size of 300 data conditions. The pattern of the average standard error bias across the data conditions for the CISE no error model and the single indicator model followed a similar pattern as found with the average parameter estimate bias.

*Statistical power and percentage of population values in confidence interval.* Table 2 also presents the average statistical power and percentage of population values in confidence interval for each level of the design factors and averaged across all the data conditions for the Type II error conditions. Statistical power values of 100.0 indicate that the null hypothesis of a population zero  $\beta_1$  path coefficient was rejected in all of the replications. With well-behaved solutions, the percentage of population values in the 95% confidence interval approximate the nominal value of 95%. Average statistical power (ASP) was highest for the CISE no error model (ASP = 85.3), followed by the CISE alpha model (ASP = 78.2), MISE model (ASP = 75.6), and the single indicator model (ASP = 59.0). Although the CISE no error model had the highest average statistical power, the average percentage of times the population value was in the 95% confidence interval (APVCI = 89.1) was lower than for the CISE alpha model (APVCI = 94.8) and the MISE model (APVCI = 94.4), but higher than for the single indicator model (APVCI = 68.9). The statistical power advantage to detect a statistically significant  $\beta_1$  path coefficient for the CISE no error model is a function of the underestimation of the standard error for the  $\beta_1$  path coefficient. Statistical power increased with decreasing residual error and number of items and with increasing reliability, sample size, and effect size for the four data analytic methods. There were only trivial differences in the average percentage of times the population value was in the 95% confidence interval across the data conditions for the MISE model and the CISE alpha model.

**TABLE 2. AVERAGE PARAMETER ESTIMATE BIAS, STANDARD ERROR BIAS, AVERAGE STATISTICAL POWER, AND NUMBER OF TIMES THE POPULATION VALUE IN 95% CONFIDENCE INTERVAL FOR TYPE II ERROR CONDITIONS**

Data Conditions	MISE Model				CISE Alpha Model				CI No Error Model				Single Indicator Model					
	APB		Power		95%		ASEB		Power		95%		APB		Power		95%	
<b>Reliability</b>	.70	102.6	93.1	67.9	94.2	100.7	98.1	71.9	95.0	81.5	78.4	82.1	85.9	50.4	51.5	46.8	56.4	87.3
	.90	100.5	97.7	87.1	94.8	100.1	98.2	87.6	94.6	94.0	95.2	89.9	93.9	79.0	79.9	77.2	87.3	
<b>Items</b>	3	101.6	94.9	77.0	94.5	100.3	98.3	78.2	94.9	86.6	85.1	85.3	89.1	66.2	66.1	66.3	72.6	
	6	101.9	94.9	74.3	94.4	100.6	97.9	78.1	94.8	88.6	85.2	85.1	89.2	57.6	59.8	51.7	65.2	
<b>Residual Error</b>	30%	101.5	95.7	80.0	94.6	100.4	98.3	82.4	95.0	88.6	85.9	90.6	89.5	65.0	62.6	67.1	68.1	
	50%	102.1	94.3	75.5	94.3	100.5	97.7	78.2	94.7	86.7	85.1	85.4	89.2	62.0	62.3	57.0	69.8	
	70%	101.7	94.8	71.4	94.4	100.5	98.4	74.0	94.8	84.4	84.4	79.7	88.6	58.7	63.9	53.0	68.7	
<b>Sample Size</b>	75	103.2	91.0	52.2	94.0	100.8	97.4	57.2	95.0	86.6	90.4	67.8	92.2	62.1	71.6	41.1	80.6	
	150	101.5	95.4	79.4	94.4	100.1	98.5	81.9	94.8	86.5	85.9	89.3	89.9	61.8	65.1	56.7	72.9	
	300	100.6	98.4	95.3	94.9	100.4	98.5	95.5	94.7	86.5	79.2	98.7	85.3	61.9	52.0	79.2	53.0	
<b>Effect</b>	5%	101.5	95.0	67.2	94.7	100.0	98.0	69.8	94.8	88.1	90.7	78.9	92.2	64.3	69.4	53.0	77.0	
	10%	102.0	94.8	84.1	94.2	100.9	98.3	86.6	94.8	85.0	79.7	91.7	86.0	59.6	56.5	65.0	60.7	
<b>Overall Mean</b>		101.8	94.9	75.6	94.4	100.5	98.1	78.2	94.8	86.6	85.2	85.3	89.1	61.9	62.9	59.0	68.9	
<b>Baseline</b>		100	95	80	95	100	95	80	95	100	95	80	95	100	95	80	95	

## 5. STUDY 2

The results of Study 1 revealed that the MISE model and the CISE alpha model outperformed the CISE no error model and the single indicator model in both Type I and Type II data conditions. The CISE alpha model generally had slightly more accurate estimates of the path coefficient and the associated standard error as well as empirical Type I error rates that slightly more closely approximated the nominal rate and slightly higher statistical power than the MISE model. Further, there were fewer convergence and offending estimate problems with the CISE alpha model than with the MISE model.

One issue with the CISE alpha model is that the reliability of the scale and the associated fixed error variance is determined based on the calculated Cronbach's alpha for the scale. As a result, the error variance for the summated scale is not simultaneously estimated with the other parameters in the model. Another approach is to fix the reliability of the scale and the associated measurement error variance to a *priori* values that might be obtained from previous research or theoretical considerations. A potential problem is that the *a priori* value for measurement error might be either an over- or under-estimate of the true measurement error. We designed a second study to explore these possibilities.

The design considerations for study 2 were the same as for study 1, except that the reliability of the composite scale was manipulated with two levels: .6 and .8. We varied the reliability to obtain greater potential differences across and within the data analytic techniques.

The following six data analytic techniques were used in Study 2: 1) MISE model in which the error variances of multiple indicators were freely estimated, 2) CISE model with error variance fixed using alpha obtained from calculated Cronbach's alpha (CISE alpha), 3) CISE model with error variance fixed using a reliability of 0, 4) CISE model with error variance fixed using a reliability of .6, 5) CISE model with error variance fixed using a reliability .8, and 6) single indicator model.

## 6. RESULTS

As with Study 1, the focus is on the  $\beta_1$  path coefficient. Bias in the estimation of the  $\beta_1$  path coefficient and standard errors associated with the estimated parameter along with Type I error rates and statistical power were calculated.

For the purposes of interpreting the results, we organized the results from the CISE models using error variance fixed to a reliability of .8 and .6 as methods either over-specifying, accurately specifying or under-specifying the true population reliability. For example, in the data conditions with a true population reliability of .8, a CISE model with error fixed using a reliability of .8 represents accurately specifying measurement error (CISE accurate), a CISE model with error fixed using a reliability of .6 represents under-specifying measurement error by .2 reliability units (CISE-.2), and CISE model with error variance fixed using a reliability of 0 represents over-specifying measurement error by .2 reliability units (CISE+.2). In the data conditions with a true population reliability of .6, a CISE model with error fixed using a reliability of .8 represents over-specifying measurement error by .2 reliability units (CISE+.2), a CISE model with error fixed using a reliability of .6 represents accurately specifying measurement error (CISE accurate), and CISE model with error variance fixed using a reliability of 0 represents over-specifying measurement error by .4 reliability units (CISE+.4).

### 6.1. Type I Error Conditions

*Offending estimates and non-converging solutions.* On average per 2,000 replications, the MISE model had 158.0 solutions with offending estimates, the CISE+.4 model had 792.1 solutions with offending estimates, the CISE alpha model had 43.3 solution with offending estimates, and the CISE accurate model had 40.8 solutions with offending estimates. The MISE model had the highest average number of non-converging solutions per 2,000 replications (10.72 solutions), followed by the CISE accurate model (0.1 solutions). The patterns across the data conditions were similar to those found in Study 1.

*Parameter estimate error and standard error bias.* Table 3 presents the average parameter estimate error and standard error bias for each level of the design factors and averaged across all the data

conditions for the Type I error conditions. Average parameter estimate bias for the  $\beta_1$  path coefficient was trivial for CISE accurate model (APB = -.002), CISE alpha model (APB = -.002), and the MISE model (APB = -.003). Over-estimation of the  $\beta_1$  path coefficient occurred for the CISE+.2 model (APB = .028), single indicator model (APB = .037), and the CISE+.4 model (APB = .040). The pattern of the APB values across the data conditions for the full model, CISE alpha model, and the CISE accurate model were similar and are consistent with the findings for the MISE model and the CISE alpha model found in Study 1.

Average standard error bias was least for the CISE accurate model (ASEB = 99.5), followed by the CISE alpha model (ASEB = 99.2), the CISE-.2 model (ASEB = 101.2), the CISE+.2 model (ASEB = 91.8), the MISE model (ASEB = 91.7), the single indicator model (ASEB = 87.7), and the CISE+.4 model (ASEB = 86.1).

*Type I error rate.* Table 3 also presents the average Type I error rate for each level of the design factors and averaged across all the data conditions for the Type I error conditions. Averaged across the data conditions, the average Type I error rate most closely approximated the nominal error rate of 5% for the CISE accurate model (4.7%), followed by the CISE alpha model (4.6%), the MISE model (4.1%), and the CISE-.2 model (3.8%). The CISE+.2 model (7.2%), the single indicator model (8.6%), and CISE+.4 model (9.2%) had severely inflated average Type I error rates.

## **6.2. Type II Error Conditions**

*Offending estimates and non-converging solutions.* On average per 2,000 replications, the CISE+.4 model had the highest number of solutions with offending estimates (714.9 solutions), followed by the MISE model (144.3 solutions), the CISE accurate model (46.3 solutions), the CISE alpha model (39.8 solutions), and the CISE+.2 model (0.7 solutions). The MISE model had the highest average number of non-converging solutions per 2,000 replications (9.78 solutions), followed by the CISE+.4 model (0.1 solutions). The patterns across the data conditions were similar to those found in Study 1.

*Parameter estimate and standard error bias.* Table 4 presents the average parameter estimate error and standard error bias for each level of the design factors and averaged across all the data conditions for the Type II error conditions. Averaged across the data conditions, average parameter estimate bias was least for CISE accurate model (APB = 100.3), followed by the CISE alpha model (APB = 101.2) and the MISE model (APB = 103.4). The CISE-.2 model severely over-estimated the  $\beta_1$  path coefficient (APB = 114.3) and the CISE+.2 model (APB = 85.3), the CISE+.4 model (APB = 70.9), and the single indicator model (APB = 50.9) severely under-estimated the  $\beta_1$  path coefficient.

Averaged across the data conditions, average standard error bias was least for the CISE accurate model (ASEB = 100.1), followed by the CISE alpha model (ASEB = 101.2), the CISE-.2 model (ASEB = 102.6), and the MISE model (ASEB = 91.6). Severely under-estimated standard errors were obtained with the CISE+.2 model (ASEB = 86.8), the CISE+.4 model (ASEB = 66.4) and the single indicator model (ASEB = 50.9). The patterns for the three accurate methods across the data conditions are similar and are consistent with the patterns found for the full method and CISE alpha model method in Study 1.

**TABLE 3. AVERAGE PARAMETER ESTIMATE ERROR, STANDARD ERROR BIAS, AND TYPE I ERROR RATE FOR TYPE I ERROR CONDITION**

Data Conditions	MISE			CISE Alpha			CISE +.4			CISE +.2			CISE Accurate			CISE -.2			Single Indicator			
	APE	ASEB	Type I Err	APE	ASEB	Type I Err	APE	ASEB	Type I Err	APE	ASEB	Type I Err	APE	ASEB	Type I Err	APE	ASEB	Type I Err	APE	ASEB	Type I Err	
<b>Reliability</b>																						
.60	.005	89.0	3.2	.003	100.8	3.7	.040	86.0	9.2	.030	93.7	6.5	.003	101.2	3.9	na	na	na	.003	89.9	7.7	
.80	.002	94.3	5.1	.002	97.6	5.4	na	na	na	.026	89.8	7.9	.001	97.9	5.4	.053	101.2	3.8	.041	85.5	9.5	
<b>Items</b>																						
3	.003	91.2	4.5	.002	98.3	4.6	.039	85.4	9.5	.028	90.9	7.4	.002	98.8	4.8	.052	101.3	3.6	.040	85.9	9.4	
6	.004	92.1	3.8	.003	100.1	4.6	.040	86.8	9.0	.028	92.6	7.0	.002	100.2	4.6	.054	101.1	4.0	.034	89.5	7.8	
<b>Residual Error</b>																						
30%	.001	91.6	4.1	.001	99.3	4.5	.048	80.4	11.5	.035	87.7	8.7	.001	99.2	4.7	.029	115.5	1.3	.046	83.3	10.5	
50%	.003	91.0	3.9	.002	99.0	4.7	.041	85.6	9.4	.028	92.1	7.1	.003	99.6	4.8	.065	95.9	4.8	.037	88.1	8.3	
70%	.007	92.5	4.4	.005	99.2	4.5	.031	92.3	6.7	.021	95.5	5.8	.003	99.8	4.6	.064	92.2	5.4	.028	91.6	7.0	
<b>Sample Size</b>																						
75	.005	84.9	3.9	.003	99.1	4.3	.041	90.4	7.5	.028	93.5	6.5	.003	99.5	4.4	.051	110.9	1.8	.036	91.4	7.0	
150	.003	93.5	4.0	.003	99.3	4.6	.039	88.1	8.2	.028	93.0	6.8	.002	99.6	4.8	.052	98.8	3.8	.038	88.7	8.2	
300	.001	96.6	4.5	.002	99.2	4.8	.040	79.8	11.9	.028	88.7	8.3	.001	99.5	4.8	.054	94.1	5.8	.036	82.9	10.6	
<b>Overall</b>																						
Mean	.003	91.7	4.1	.002	99.2	4.6	.040	86.1	9.2	.028	91.8	7.2	.002	99.5	4.7	.053	101.2	3.8	.037	87.7	8.6	
Baseline	0	95	5	0	95	5	0	95	5	0	95	5	0	95	5	0	95	5	0	95	5	

**TABLE 4. AVERAGE PARAMETER ESTIMATE BIAS, STANDARD ERROR BIAS, AVERAGE STATISTICAL POWER, AND NUMBER OF TIMES THE POPULATION VALUE IN 95% CONFIDENCE INTERVAL FOR TYPE II ERROR CONDITIONS**

Data Conditions	CISE Accurate				CISE -.2				Single Indicator			
	APB	SEB	Pow	95%	APB	SEB	Pow	95%	APB	SEB	Pow	95%
<u>Reliability</u>												
.60	99.9	102.2	52.8	96.2	na	na	na	na	36.9	43.8	28.3	42.2
.80	100.6	97.9	79.9	94.6	114.4	102.6	62.7	95.4	60.2	58.0	57.7	65.2
<u>Items</u>												
3	100.7	100.8	66.7	95.5	114.9	103.6	62.6	95.4	57.2	56.7	53.5	62.3
6	99.9	99.4	66.0	95.2	113.8	101.6	62.8	95.3	39.9	45.2	32.5	45.1
<u>Residual Error</u>												
30%	99.4	100.3	69.3	95.2	104.3	117.8	62.5	98.1	51.0	51.6	47.4	54.8
50%	100.1	100.2	66.7	95.6	113.8	100.4	62.3	95.3	48.5	51.8	42.1	54.7
70%	101.3	99.7	63.1	95.3	124.3	89.8	63.3	92.6	46.1	49.3	39.5	51.6
<u>Sample Size</u>												
75	99.9	101.0	42.3	95.9	113.6	110.8	34.5	97.7	48.4	62.0	25.3	72.5
150	100.7	98.1	68.1	94.8	114.6	100.2	65.5	94.9	49.2	52.2	41.2	56.5
300	100.2	101.1	88.5	95.4	114.8	96.9	87.9	93.4	48.1	38.6	62.5	32.1
<u>Effect</u>												
5%	99.9	100.1	56.4	95.5	109.9	107.3	49.1	96.7	51.4	59.3	36.8	67.1
10%	100.6	100.0	76.3	95.2	118.8	97.9	76.3	94.0	45.7	42.5	49.2	40.3
Overall Mean	100.3	100.1	66.3	95.4	114.3	102.6	62.7	95.4	48.6	50.9	43.0	53.7
Baseline	100	95	80	95	100	95	80	95	100	95	80	95

*Statistical power and percentage of population values in confidence interval.* Table 4 also presents the average statistical power and percentage of population values in 95% confidence interval for each level of the design factors and averaged across all the data conditions for the Type II error conditions. As with Study 1, statistical power values of 100.0 indicate that the null hypothesis of a population zero  $\beta_1$  path coefficient was rejected in all of the replications and well behaved percentage of population values in the 95% confidence interval approximate the nominal value of 95%.

Averaged across the data conditions, the highest average statistical power was obtained with the CISE+.2 model (ASP = 75.8), followed by the CISE+.4 model (ASP = 71.2), the CISE accurate model (ASP = 66.3), the CISE alpha model (ASP = 66.0), the CISE-.2 model (ASP = 62.7), the MISE model (ASP = 61.1), and the single indicator model (ASP = 43.0). As with Study 1, the methods that did not perform well estimating the  $\beta_1$  path coefficient demonstrated inflated power because these methods severely underestimated the standard errors, resulting in upwardly biased *t*-values.

The percentage of population values in 95% confidence interval index closely approximated the nominal value of 95% for the CISE alpha model (95.2), CISE accurate model (95.4), the CISE-.2 model (95.4), and the MISE model (94.1). Although the CISE-.2 model performed well on average, the method did not perform well in some of the data conditions such as the sample size of 75 and 300 and the small and large residual error data conditions. On the other hand, the CISE alpha model, CISE accurate model, and the MISE model had values that reasonably approximated the nominal value across all the data

conditions. The percentage of population values in 95% confidence interval index were unacceptably low for the CISE+.2 model (90.3), the CISE+.4 model (76.6), and the single indicator model (53.7).

*Summary.* Overall, the CISE accurate model and the CISE alpha model performed the best across the data conditions for both the Type I and Type II error conditions, followed by the MISE model. Inaccurate CISE models were problematic.

## 7. DISCUSSION

Methods for handling measurement error in regression analysis range from assuming that measures are perfectly reliable to explicitly modeling measurement error. In our two simulation studies, we examined the effectiveness of strategies for incorporating measurement error in regressions. The worst performing strategy was to assume that each construct in the regression model can be represented by a single infallible indicator. Biased estimated parameters and standard errors, inflated Type I errors, and lower statistical power were yielded when a single item with no corrections for measurement error represented each construct in the regression analysis. The recognition that single unreliable indicators of a constructs in regression models perform poorly certainly is not a new revelation and is widely recognized in most research methods books that deal with regression (e.g., Cohen & Cohen, 1983, Pedhazur & Schmelkin, 1991). Our results reinforce this notion and reveal the pattern of biases in parameter estimates and standard errors.

In response to the widespread recognition that single indicators of constructs are usually unreliable, many researchers use a composite scale comprised of more than one item. The assumption is that measurement errors are not systematic and unreliability in one item will be offset by unreliability in other items. Unfortunately combining items into a composite scale does not eliminate all measurement error. In this study, the use of a composites with no further estimation and specification of measurement error suffered, although less severely, the same problems as with using fallible single indicators in a regression model: inflated Type I errors and reduction in statistical power.

With the emergence of structural equation modeling, researchers are able to explicitly model measurement error in regression models in a variety of ways. One way to explicitly model measurement error is to represent a latent construct with several indicators in a multiple indicator structural equation model (MISE). As long as either one of the paths from a latent construct to the associated indicators is fixed or the variance of the latent construct is fixed, measurement error can be simultaneously estimated with the parameters in the structural portion of the model (i.e., relations among the latent variables). Our results revealed that the MISE method performed reasonably well in estimating parameter estimates, but tended on average to under-estimate standard errors, especially with small samples and unreliable indicators. This finding is not overly surprising because this method relies on maximum likelihood estimation that is based on asymptotic assumptions. Our results revealed that with samples of 300, the bias in standard errors was approximately the same as the method that most accurately estimated the standard error (i.e., the CISE alpha method). A potential problem with the MISE method is that convergence and offending estimates can be troublesome with small samples and complex models when a large number of parameters are estimated (Tanaka, 1987). Although our model used in the two studies was not overly complex, non-converging solutions and offending estimates were encountered, especially with data conditions characterized by unreliable indicators and small sample sizes.

Another way to explicitly model measurement error is to fix the unreliability and associated measurement error of an indicator to a specific value. This approach is usually combined with using composite scales as indicators of related constructs in a composite structural equation modeling method (CISE alpha model). By combining the indicators of a construct into a composite scale, reliability can be estimated from the formula for Cronbach's alpha and the reliability of the indicator in the structural equation model can be fixed to this value. The main advantage of this approach is that the number of estimated parameters is reduced because the complexity of the structural equation model is reduced with the simplification of the measurement portion of the overall model. A concern with this approach is that measurement error and the structural parameters are not simultaneously estimated. The results of our two simulation studies revealed that the CISE model with alpha was the most effective method in

estimating parameter estimates and standard errors as well as controlling the Type I error rate and maximizing statistical power across the data conditions. Further, there was fewer convergence and offending estimate problems with the CISE alpha model than with the MISE model. This supports the notion that reducing complexity and estimation demands in a structural equation model reduces estimation problems such as non-converging solutions and offending estimates.

Some may criticize fixing measurement error and indicator reliabilities in a structural equation model to values estimated from the same data, but not estimated simultaneously with the structural parameters. To overcome this potential problem, indicator reliabilities and associated calculated measurement error can be fixed to *a priori* values, that are based on past research, theory, pilot studies or holdout samples. This approach is in the spirit of errors-in-variables regression (Heise, 1975, 1986; Fuller & Hidioglou, 1978; Warren, Keller-White, & Fuller, 1974). The results of second study revealed that this approach performs well only when the *a priori* values approximate the true population values. When *a priori* values that are not representative of the true population values are used to fix indicator reliabilities and measurement error, bias in parameter estimates and standard errors occur. When the fixed indicator reliabilities are low, the resulting fixed measurement errors are too large. This tends to yield upwardly biased parameter estimates and standard errors. Downwardly biased parameter estimates and standard errors are encountered when fixed reliabilities are too large and the associated fixed measurement error is too small.

In sum, the CISE alpha model performed the best on average across the data conditions. With large samples and reliable indicators (i.e., items), the MISE model method is approximately as effective as the CISE alpha model method; this is not surprising because MISE is a maximum likelihood estimation method based on asymptotic assumptions. However, it is important that, when using the CISE method, that measurement error is accurately estimated, either through alpha or through accurate *a priori* information.

### **7.1. Limitations and Future Research**

Some limitations of the study warrant attention. Although care was taken to ensure that the data conditions used in the simulations were representative of conditions often encountered by organizational researchers, additional replications are required before one can generalize to a wider variety of situations. Second, this study focused on a relatively simple model involving only three exogenous variables and one endogenous variable. Organizational researchers often test more complex models; future research can extend these findings to a wider variety of structural models.

One area for future research concerns the issue of model mis-specification. There has been some research on this topic with item parceling (e.g., Hall, Snell, & Foust, 1999). With using a single composite scale as an indicator of a latent variable instead of multi-indicators for a latent variable, the bias that results from mis-specifying a single indicator (i.e., associating the indicator with the wrong latent variable) will be diluted when this indicator is combined with other properly specified indicators in a composite scale and the specification error in that particular composite indicator will not influence the other composite indicators. With a MISE approach, the bias from mis-specifying a single indicator will reverberate throughout both the measurement and structural portions of the model. A disadvantage of a composite structural equation modeling approach is that residuals and modification indexes for the measurement model cannot be inspected from point of stress and mis-specification because the measurement portion contains only fixed parameters (i.e., reliabilities and measurement errors). Bandalos and Finney (2001) further argued that Type II errors may be increased with indicator parceling. With a multiple indicator structural equation modeling approach, mis-specification can be identified from inspection of residuals and modification indexes.

### **7.2. Recommendations to Researchers**

Although the results of this study suggest that a single indicator composite model with measurement error fixed using reliabilities estimated from the formula for Cronbach's alpha performs well across a variety of data conditions, the choice of the CISE alpha model or the MISE model should depend on the research questions being investigated and the psychometric history of the measures. Bagozzi and Edwards (1998) presented a general approach for representing constructs in organizational research that should



help guide researchers in model choices. If a researcher conceptualizes the study constructs as global phenomenon, then a single indicator composite model is appropriate. However, an atomistic conceptualization of the constructs calls for a multiple indicator structural equation modeling approach that requires larger sample sizes.

The psychometric history of the items and scales used to measure the constructs will also affect the choice of the type of model specification. A single indicator composite model should only be used if the items and scales used to measure the constructs have well-established psychometric properties and the composite scales are unidimensional. Because residuals and modification indexes are not provided with a single composite indicator structural equation model, mis-specification in the relationships between constructs and indicators cannot be identified. As a result, single composite indicator structural equation models can mask mis-specification such as secondary influences (Bandalos & Finney, 2001; Hall, Snell, & Foust, 1999). Bagozzi and Edwards (1998) noted that "it is important that items are developed carefully and that exploratory factor analyses at the individual item level support any aggregation into indicators" (p. 55). Unlike the composite indicator structural equation modeling approach, a multiple indicator structural equation modeling approach allows for a researchers to examine measurement issues, but larger samples are required than with a single composite indicator structural equation model.

A review of organizational research reveals that many researchers have already adopted the use of CISE alpha models (e.g., Frone, Yardley & Markel, 1997). However, to our knowledge, this is the first study to empirically demonstrate that this method for dealing with measurement error is valid, as long as the researcher is aware of the potential limitations. When appropriately applied, the CISE alpha method can result in accurate structural equation models. Researchers need to be aware of both the advantages and disadvantages of using this method, and to choose the method for handling measurement error that is most appropriate for their purposes.

#### REFERENCES:

- Anderson, L.E., Stone-Romero, E.F., & Tisak, J. (1996). A comparison of bias and mean squared error in parameter estimates of interaction effects: Moderated multiple regression versus errors-in-variables regression. *Multivariate Behavioral Research*, *31*, 69-94.
- Babakus, E., Ferguson, C.E., & Joröskog, K.G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, *24*, 222-228.
- Bagozzi, R.P., & Edwards, J.R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, *1*, 45-87.
- Bandalos, D.L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, *9*, 78-102.
- Bandalos, D.L., & Finney, S.J. (2001). Item parceling issues in structural equation modeling. In G.A. Marcoulides & R.E. Schumacker (Eds.), *New developments and techniques in structural equation* (pp. 269-296). Mahawah, NJ: Lawrence Erlbaum Associates, Inc.
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: John Wiley.
- Bollen, K.A. (1996). A limited-information estimator for LISREL models with or without heteroscedastic errors. In G.A. Marcoulides and R.E. Schumacker (eds.), *Advanced Structural Equation Modeling: Issues and Techniques*. Lawrence Erlbaum Associates: Mahwah, NJ.
- Cohen, J. & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling*, *9*, 327-346.
- Finch, J.F., West, S.G., & MacKinnon, D.P. (1997). Effects of sample size and nonnormality on the estimation of mediated effects in latent variable models. *Structural Equation Modeling*, *4*, 87-107.
- Frone, M.R., Yardley, J.K. & Markel, K.S. (1997). Developing and testing an integrative model of the work-family interface. *Journal of Vocational Behavior*, *50*, 145-167.
- Fuller, W.A., & Hidiroglou, M.A. (1978). Regression estimation after correcting for attenuation. *Journal of American Statistical Association*, *73*, 99-104.

- Green, S.B., Akey, T.M., Flemming, K.K., Hersberger, S.L., & Marquis, J.G. (1997). Effect of the number of scale points on chi-square fit indices in confirmatory factor analysis. *Structural Equation Modeling, 4*, 108-120.
- Hall, R.J., Snell, A.F., & Foust, M.S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary constructs. *Organizational Research Methods, 2*, 233-256.
- Hayduk, L.A. (1987). *Structural equation modeling with LISREL*. Baltimore: John Hopkins University Press.
- Heise, D.R. (1975). *Causal analysis*. New York, NY: Wiley.
- Heise, D.R. (1986). Estimating nonlinear models: Correcting for measurement error. *Sociological Methods & Research, 14*, 447-472.
- Hinkin, T.R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*, 967-988.
- Huchthinson, S.R., & Olmos, A. (1998). The behavior of ad hoc fit indices in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling, 8*, 115-130.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL VIII user's reference guide*. Chicago, IL: Scientific Software International.
- Landis, R.S., Beal, D.J., & Tesluk, P.E. (2000). A comparison of approaches to forming composite measures in structural equation models. *Organizational Research Methods, 3*, 186-207.
- Little, T.D., Cunningham, W.A., Shahar, G., & Widaman, K.F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151-173.
- Muthen, B.O., & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38*, 171-189.
- Pedhazur, E.J., & Schmelkin, L.P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rushton, J.P., Brainerd, C.J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin, 94*, 18-38.
- Russell, C.J., & Bobko, P. (1992). Moderated regression analysis and Likert scales: Too coarse for comfort. *Journal of Applied Psychology, 77*, 336-342.
- Russell, C.J., Pinto, J.K. & Bobko, P. (1991). Appropriate moderated regression and inappropriate research strategy: A demonstration of information loss due to scale coarseness. *Applied Psychological Measurement, 15*, 257-266.
- Tanaka, J.S. (1987). "How big is enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development, 58*, 134-146.
- Warren, R.D., Keller-White, J., & Fuller, W.A. (1974). An errors-in-variables analysis of managerial role performance. *Journal of American Statistical Association, 69*, 886-893.
- Xitao, F., Thompson, B., & Wang, L. (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling, 6*, 56-83.