Factor Analysis in Educational Settings: A Simulation Study Comparing Fit Statistics Across Robust Estimators

R. Noah Padgett, M.A.

Chairperson: Grant B. Morgan, Ph.D.

In education and social science, data often arise from nested data structures, meaning that students are nested within teachers or schools. Traditional factor analytic approaches to measuring latent traits do not account for the nested structure of these data. The logic and potential issues of using multilevel confirmatory factor analysis were discussed. The ability of commonly used fit statistics to discriminate between a correctly specified model and models with omitted factor loading(s) were investigated with receiver-operating-characteristics (ROC) analyses. Combining ROC analyses with traditional methods of investigating fit statistic performance resulted in converging evidence for the utility of these common fit statistics. In general, these fit statistics performed poorly and should not be heavily relied upon for evidence of the factor structures specified. Recommendations were given for which commonly reported fit statistics to use, cut-off criteria to use for which estimators, and cautions about the use of the suggested cut-off criteria.

Factor Analysis in Educational Settings: A Simulation Study Comparing Fit Statistics Across Robust Estimators

by

R. Noah Padgett, B.S.

A Thesis

Approved by the Department of Educational Psychology

Grant B. Morgan, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Master of Arts

Approved by the Thesis Committee

Grant B. Morgan, Ph.D., Chairperson

Nicholas F. Benson, Ph.D., NCSP

Amanda S. Hering, Ph.D.

Accepted by the Graduate School
August 2019

J. Larry Lyon, Ph.D., Dean

*Page bearing signatures is kept on file in the Graduate School.*

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

| | |
|---|---|
| CFA | Confirmatory Factor Analysis |
| MCFA | Multilevel Confirmatory Factor Analysis |
| SEM | Structural Equation Modeling |
| ANOVA | Analysis of Variance |
| ROC | Receiver Operating Characteristic |
| AUC | Area Under the Curve |
| MLR | Maximum likelihood with robust standard errors |
| USLMV | Unweighted least squares mean and variance adjusted |
| WSLMV | Weighted least squares mean and variance adjusted |
| CFI | Comparative Fit Index |
| TLI | Tucker-Lewis Index |
| RMSEA | Root Mean Square Error of Approximation |
| SRMRW | Standardized Root Mean Residual Within |
| SRMRB | Standardized Root Mean Residual Between |
| $N_1$ | Number of units sampled at level-1 per group |
| $N_2$ | Number of groups sample at level-2 |
| $ICC_O$ | Observed variable Intraclass correlation coefficient |
| $ICC_L$ | Latent variable Intraclass correlation coefficient |

# CHAPTER ONE

## Introduction

Factor analysis models have been utilized in educational and psychological investigations for over a 100 years to help measure constructs that we cannot directly or easily measure. Factor analysis provides a flexible framework for measuring traits and constructs. However, in educational settings some constructs cannot be studied fully within the traditional factor analytic framework because students may be influenced by factors that are contextual or beyond the students' control. For example, studying how students perceive the climate at their school is necessarily influenced by factors that directly affect students and factors that influence the environment in which students exist (e.g., school). When measuring outcomes from students both levels of influence need to be accounted for in order to gain a full understanding for how students respond on, say, psychological assessments. In this context, one of the most commonly employed data collection methods is a self-report measure where students indicate the degree to which they agree or disagree to a set of statements. However, the data that arise from this form of assessment in an educational setting have multiple layers of complexity that need to be accounted for in order to more accurately reflect the processes that influence responses.

### *Purpose of Study*

In this study, the effects of ordered categorical data on the ability of commonly used fit statistics to detect incorrectly specified measurement models was examined. Given that data from educational settings typically arise from hierarchically structured data and commonly use ordered categorical response formats, there is a need to understand how these models can be examined. Establishing the adequacy with which a model explains the interrelationships among items (i.e., measurement model) is therefore a central issue that needs to be considered in these complex yet common scenarios. Due to the limitations of prior studies of the available fit statistics for multilevel measurement models, the outcome of this study is to determine whether specific cut-off values are indicative of global misfit of the level-1 and/or level-2 model. Therefore, the overall purpose is to provide practitioners with the knowledge of how fit statistics work in helping identify model fit and what values they need to consider for these determinations.

### *Overview of Procedures*

In this Monte Carlo simulation study, a fully crossed design among factors with 500 replications per cell of the design was used. The factors varied across conditions were: three levels of observed variable intra-class correlation coefficient, two levels of latent variable ICC, three different sample sizes within group, and four different number of groups sampled. This design yielded a total of 72 ($3 \times 2 \times 3 \times 4$) unique data conditions. For each of the datasets generated, two factors were crossed for examination estimation: three estimators and fours type of model specification. In multilevel confirmatory factor analysis, the measurement model can be misspecified at level-1 and/or level-2. Because of the complexity of specifying these models, the four specifications align with four different ways of correctly or incorrectly specifying the entire model. This yielded 12 ($3 \times 4$) models estimated for each generated dataset. Convergence rates of models across conditions and estimators were also examined.

For each replication across conditions and across estimators, the ability of commonly used fit statistics to identify misfit was examined. Using receiver-operating-curve (ROC) analyses, the performance of common fit statistics was assessed by systematically varying the cut-off criteria for determining how a model fits. The ROC analyses estimated area under the curve (AUC), and the hit rates were used to determine what values of each fit statistic help differentiate between a fitting and non-fitting model. For these analyses, the aim was to help determine if these fit statistics could identify the correctly specified model versus models that have incorrectly specified measurement models at either level and whether certain fit statistics were sensitive to only certain type of misfit.

### *Delimitations*

In this study, the small foundation is extended for examining multilevel measurement models with categorical data, specifically with regard to how common fit statistic performance across different estimators for categorical data. As with any simulation study, the results of only generalize to the limited conditions examined. The conditions chosen were selected to mirror conditions of applied researchers as close as possible while still being as parsimonious as possible. The restriction of conditions was due partially to how long data generation took and model estimation took. Estimation with MLR was observed to result in most usable cases per cell on average, but this may have occurred because data were treated as continuous. Using MLR in M*plus* with categorical data requires numerical integration across four dimensions (one dimension per latent variable) that is computationally burdensome, so I assumed these data were continuous in order to estimate these models in a reasonable amount of time. Additionally, the cells of this design ended up with unequal sample sizes due to convergence issues and the number usable replications. Even in some cells, the number of usable cases was zero. These factors that limit the sample size in some cells could have influenced these results in unaccounted for ways and caused the impact of some design factors to be underestimated. Future work could address how to estimate these types of models under conditions that failed to converge or provide useful information.

The use of only one type of model misspecification (i.e., an omitted crossloading) limits the generalizability of these finding. Other potential sources of misspecification (i.e., residual correlations among items, omitted factor correlation(s), wrong number of factor(s) specified, etc.) may result in different conclusions about distributions of these fit statistics. Other types of misspecification can be investigated with these data, and the generated data are available online (Padgett, 2019).

# CHAPTER TWO

## Literature Review

In this chapter, I review the pertinent parts of the literature on multilevel modeling, confirmatory factor analysis (CFA), categorical CFA, multilevel CFA, estimation methods, and fit statistics.

### *Multilevel Modeling*

Data often arise from hierarchical and organizational settings. For example, a school is organized hierarchically such that children are grouped into classrooms, and a school itself is organized hierarchically within a district, and so on. When data on students are obtained across classrooms, schools, and/or districts, these data are called hierarchically structured or have a nested structure. When data arise from multiple levels of sampling, variables are therefore measured at different levels. In education, student characteristics are typically the lowest level of analysis, called level-1. Characteristics that reflect the classroom are the next level of analysis, level-2. As the lowest level of dat collection, the characteristics may refer to the *micro level*, whereas data from higher levels are *macro level* (Heck & Thomas, 2015). Macro level variables refer to contexts or groups because each level of a macro level variable will subsume multiple cases from a lower-level.

As an example, the nested data structure of educational data across schools creates difficulties in investigating student dropout factors. Many factors can be observed that as associated with dropout such race, SES, family, etc. However, other factors may also influence why dropout rates vary among schools. Rumberger (1995) discussed the difficulties is identifying the influencers of student dropout because groups of students will have the same school factors acting on the potential for dropout.

He noted how existing literature on investigating dropout has focused on individual factors, such as family background, early school experiences, school attendance, behavior, and other demographic factors. However, Rumberger used this research to investigate school and community factors that influence dropout rates. Student composition is an example of a broad range of factors that influence the effectiveness of schools. Furthermore, school climate characteristics are measured by school-level variables that potentially influence the institutional dropout rate. The fact that students were sampled within the same school means that the same factors influence each case, creating dependence among the cases.

When students are assessed over time, observations are potentially doubly nested. In this example, each student has multiple measurement occasions, and students may be nested within classrooms. The measurements are nested with the student, and the students are nested within classrooms (Aitkin & Longford, 1986). This scenario would be an example of a three-level data structure. Other examples of data with a hierarchical structure are

(1) longitudinal data, in which individuals have repeated measurements of the same outcome, and

(2) Meta-analysis, in which effect sizes are drawn from a sample of studies.

Interested readers are referred to Raudenbush and Bryk (2002), Heck and Thomas (2015), and Hox, Moerbeek, and Van de Schoot (2017) for more examples.

The analysis of data arising from such scenarios must account for the hierarchical structure in order to gain insight into student growth or organizational effects. Traditional methods of analysis such as ordinary least squares regression, fail to account for the nested structure (Raudenbush & Bryk, 2002). In educational settings, hierarchical linear models have been used to account for the nested structure. For example, Bryk and Driscoll (1988) investigated teachers' self-efficacy across schools. An average of 22 teachers were measured across 375 schools, and one of the main goals of the analysis was to investigate the effects of school characteristics on teacher's self-efficacy. Ignoring the multilevel structure of these data typically leads to underestimated fixed-effect parameters and standard errors, which can lead to inflated Type-I error rates(Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). These effects are due to a violation of independence in the distribution of error. In simple linear regression, the assumption is that cases are independent, and multilevel modeling aims to account for the dependence between errors associated with violating this independence. Note that much of this literature is related to multilevel linear models, which is the extension of linear regression to multilevel settings.

The effect of ignoring the multilevel structure generally depends on the particular scenario; however, a general issue that occurs is that standard errors are underestimated (Raudenbush & Bryk, 2002; Snijders & Bosker, 1993; Van der Leeden & Busing, 1994). For example, assigning everyone in a particular group the same value of a higher level (e.g, assign all teachers to the same school the value of school size) reduces the variability in the predictor, which causes the estimation of the standard error to be downwardly biased. The amount of bias depends, in part, on the effect of group membership. The effect of group membership can be estimated by the *intra-class correlation* coefficient (ICC), which is also called the *cluster effect* (Raudenbush & Bryk, 2002). The ICC is the proportion of variance in an outcome attributed to group membership relative to the total variation in the outcome.

Investigations of the effect of varying levels of ICC on parameter estimation has often been paired with sample size evaluations. For example, sample size requirements for sufficient parameter estimation depends in part on the size of the ICC (Maas & Hox, 2005; Scherbaum & Ferreter, 2009). The necessary sample also depends on the parameter(s) of interest in the research. For example, if one is interested in the variance of level-2 parameters, especially slopes, 100 groups may not even be sufficient for adequate precision (Maas & Hox, 2005). This is in contrast to interest in the intercepts, where precision tends to be greater and sample size requirements tend to lower (Hofmann, 1997). Sample size has been studied extensively for multilevel linear models, and interested readers are referred to Bell, Morgan, Schoeneberger, Kromrey, and Ferron (2014) for a more recent examination and overview of sample size requirements in multilevel linear models.

In education, outcomes of interest are often latent, meaning that the construct is not directly observed, but rather is inferred from a set of indicators. Next, I will introduce the methods used for measuring indirectly observed constructs and will incorporate multilevel modeling of these outcomes.

### *Confirmatory Factor Analysis*

Education and social scientists have studied constructs that are not directly measurable for over a century. For example, Spearman (1904) worked on the measurement of general intelligence, a construct that is theorized to exist but cannot be directly measured. The method for relating observed indicators to continuous latent variables, such as intelligence, is known as factor analysis. Factor analysis is based on the common factor model, and the common factor model is a general model for how the relationships among observed variables are explained by unobserved, latent variables (i.e., factors).

If one were interested in measuring the length of an object she may use a ruler as an instrument to determine how many centimeters long the object is. In education, researchers are frequently interesting in measuring properties that cannot be directly measured, as was the case with the ruler. Instead, one uses tests or surveys as instruments to gain insight into the variable on interest by examining the responses to items on the instrument. The measurement process is akin to trying to identify where someone is on a ruler when the ruler is inside a box covered in cloth. If one shines a flashlight on one of the sides of the box, she can see the reflection of the ruler and where the person is on the ruler, but, because of the covering on the box, she doesn't have a perfect reflection. This is similar to asking a single item on a survey; that is, responses are related to the construct but not perfectly. Therefore, multiple items are include (i.e. shine multiple lights on the box from multiple angles) in order to triangulate the position the person is on the construct of interest. The general idea behind factor analysis is to use multiple items to triangulate the measurement of the construct.

Factors are triangulated investigating how responses to items covary. When item covary, we aim to explain why they covary by their relationship with the underlying construct (i.e., factor of interest). When responses covary, the responses to one item are related to responses to another item. Based on how the items are theorized to group together, we then can form an expectation of the level of covariance among the items. The factor model we hypothesize is then tested to see if we can explain the covariance among the set of items measured. This is why factor analysis, specifically confirmatory factor analysis (CFA), is also sometimes called covariance structure analysis (Bollen, 1989; Brown, 2015; Kline, 2016).

Over the years, a vast methodological literature has amassed on CFA. Much of this literature is out of scope for this review. However, some seminal pieces are highlighted. In the development of CFA, Jöreskog (1967) helped progress the estimation of CFA under maximum likelihood. Later, Jöreskog (1969) helped build the groundwork for hypothesis testing in CFA and SEM more generally. Possibly the greatest contribution to the CFA literature was the development of the software program LISREL by Jöreskog and Sörbom, where they provided the technical capabilities to estimate a wide range of latent variable models (Jöreskog & Sörbom, 2015). The entirety of the work of Jöreskog and colleagues is out of scope for this review, but much of the future work on CFA rests of shoulders of these giants.

Much of the work on CFA and SEM was brought together by Bollen (1989), whose text contains the synopsis of much of this early work on latent variable modeling. The general modeling framework of CFA is encapsulated in the following concise model, known as the common factor model:

$$\mathbf{Y}_i = \tau + \mathbf{\Lambda}\eta_i + \varepsilon_i \tag{2.1}$$

where $\mathbf{Y}_i$ is the vector of observed item responses of individual $i$, $\tau$ is the vector of intercepts that is typically fixed to a zero vector because of the use of standardized scores in the estimation, $\mathbf{\Lambda}$ is the factor loading matrix, $\eta_i$ is the vector of factor scores for individual $i$, and $\varepsilon_i$ is the residual error. Equation 2.1 relates the model parameters to the observed responses. However, CFA is a method for modeling how patterns of responses can be explained based on the theorized variables. The theorized relationship among variables is often expressed through a path model. An example of a single one factor CFA path model is shown in Figure 2.1.

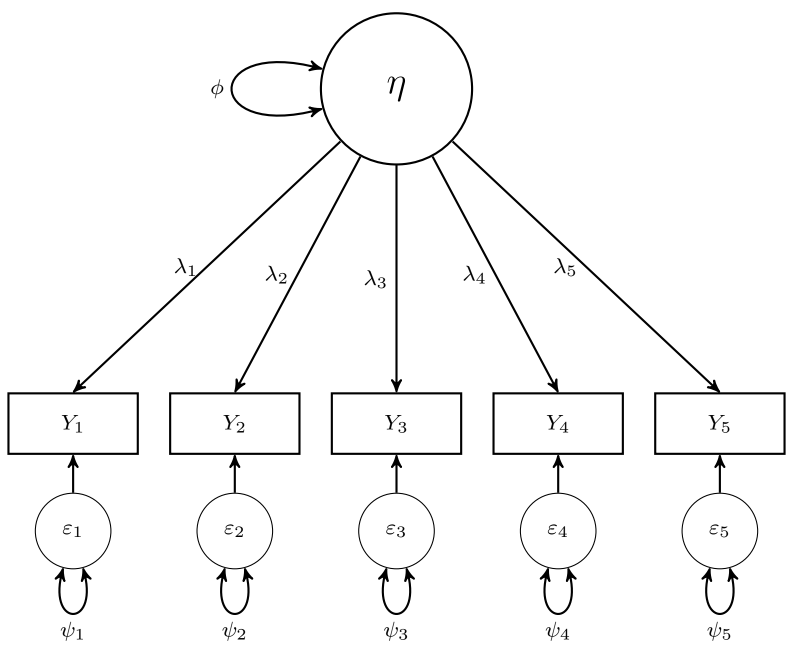



*Figure 2.1.* Example path diagram of a confirmatory factor analysis model.
*Note.* The error terms ($\varepsilon_i$) and corresponding residual variances ($\psi_i$) are typically excluded for simplicity.

Along with relating the observed responses to the latent variables, we need to specify how the variables relate to each other. Our expectation for the relationships among item is formulated as

$$Var(\mathbf{Y}) = \mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Psi}\mathbf{\Lambda}^{\mathrm{T}} + \mathbf{\Theta} \tag{2.2}$$

where $Var(\mathbf{Y})$ is the model implied covariance matrix which is often shortened to $\mathbf{\Sigma}$. And, where $\mathbf{\Lambda}$ is the estimated factor loading matrix which is sometimes called the pattern matrix, $\mathbf{\Psi}$ is the estimated covariance/correlation matrix among the latent variables, $\mathbf{\Lambda}^{\mathrm{T}}$ is the transpose of the factor loading matrix, and $\mathbf{\Theta}$ is the residual covariance matrix among the observed variables. The residual covariance matrix is typically assumed to be a diagonal matrix, which means that once the factor structure imposed on these data is accounted for no other relationship among items exists (Brown, 2015). The interested reader is referred to Bollen (1989); Brown (2015); Kline (2016) for more information on CFA.

### *Categorical CFA*

Traditional confirmatory factory analysis assumes that factor indicators are continuous measures and that these indicators are linearly related to the underlying factors. A continuous measure can (usually) be sufficiently described by its mean and variance, whereas data with discrete categories typically are not appropriately described by its mean and variance. Furthermore, categorical data are unlikely to have a linear relationship with the underlying factor(s). Social scientists rarely obtain data that are continuous in nature as surveys often use ordered response scales (e.g., Likert-type responses). CFA models fit with these types of data are sometimes called *item factor analysis* and are closely related to some item response theory models (Bollen, Bauer, Christ, & Edwards, 2010). The use of traditional CFA, particularly in conjunction with maximum likelihood methods, may not be valid given the restrictions of categorical data. Methods for circumventing the limitations of categorical data are continually being developed and investigated.

Categorical data have long been used to approximate some underlying dimension. For example, Likert-type responses of *Strongly Disagree*, *Disagree*, *Neutral*, *Agree*, and *Strongly Agree* can be seen as attempting to measure a continuous attitude of agreement towards a topic. Using discrete categories is a simple and straightforward method for measuring attitudes, and the use of this technique is likely to continue in the future. Given the prevalence and nearly certain continual use of such response formats, having methods that account for the limited range of response in the analysis is important for gaining useful information from these data.

The traditional CFA model only reflects a linear relationship between the observed scores and the factors. This implies that the observed score on an item is a linear function of the factor loading, factor score, intercept, and error. Breaking this relationship down to the item level helps with the discussion of the theory for categorical CFA. The item level relationships can be represented as follows for an item in a one factor model

$$y_{ij} = \tau_j + \lambda_j \eta_i + \varepsilon_{ij} \tag{2.3}$$

where $y_{ij}$ is the observed score of the $i^{th}$ individual on the $j^{th}$ item, $\tau_j$ is the intercept of the $j^{th}$ item, $\lambda_j$ is the factor loading for the $j^{th}$ item, $\eta_i$ is the amount that individual $i$ has on the factor, and $\varepsilon_{ij}$ is the error in measurement for the $i^{th}$ individual on the $j^{th}$ item; see Benson and Nasser (1998) for a discussion on the linear combination of these components. This linear combination is analogous to how simple linear regression relates a predictor (x or $\eta_i$) to an outcome (y). Just as ordinary least squares regression is inappropriate for categorical outcomes, such as binary outcomes, traditional CFA is also inappropriate for categorical outcomes. In place of OLS for binary or categorical outcomes, logistic regression is appropriate as the model is transformed to account for the discrete nature of the outcome. A similar idea is carried over to factor analysis where the observed scores are nonlinearly related to the factors.

The nonlinear relation is only defined for the observed categorical response. As noted earlier, the categorical responses are a coarse measurement of a continuous underlying response. The underlying continuous response is known as the *latent response distribution* (Muthén, 1984). By formulating the observed scores to be a discretized approximation to the latent response distribution, we are able to capture a linear relationship between the latent response and the factors.

$$y_{ij}^{*} = \tau_j + \lambda_j \eta_i + \varepsilon_{ij} \tag{2.4}$$

where $y_{ij}^{*}$ is the latent response score of the $i^{th}$ individual on the $j^{th}$ item, $\tau_j$ is the intercept of the $j^{th}$ item which is usually equal to zero, and the other remaining parameter are the same as before. The latent response distribution is typically a standard normal variable with a mean of zero and standard deviation of one, and the intercept is therefore zero. The burden then falls on defining and estimating the mechanism for how the observed score (y) relates to the latent response score ($y^{*}$).

The latent response score is typically defined to relate to the observed score through a simple categorization scheme. Suppose an item had three response categories coded *Disagree*, *Neutral*, *Agree*. Creating three categories from a continuous latent response requires two thresholds for categorizing. The first threshold defines whether the observed response is Disagree versus Neutral. This means that if the latent response score is lower than the value of threshold one then the observed score is Disagree. The second threshold defines the level of the latent response necessary to response Neutral versus Agree. If the latent response is greater than the first threshold but lower than second then the observed response is Neutral. A response of Agree occurs when the latent response score is greater than the second threshold. Therefore, for $C$ categories, $C-1$ thresholds are needed to relate the latent response to the observed score. Muthén (1984) expressed this relationship mathematically as system of equations

$$y_{ij} = \begin{cases} C_j - 1, & if\ \tau_{j,C_j-1} < y^*_{ij} \\ C_j - 2, & if\ \tau_{j,C_j-2} < y^*_{ij} \leq \tau_{j,C_j-1} \\ \vdots & \\ 1, & if\ \tau_{j,C_1} < y^*_{ij} \leq \tau_{j,C_2} \\ 0, & if\ y^*_{ij} \leq \tau_{j,C_1} \end{cases} \tag{2.5}$$

where $y_{ij}$ is the observed score of the $i^{th}$ individual to the $j^{th}$ item which contains $C_j$ response categories, and $C_j - 1$ category thresholds. Other authors, e.g. Dolan (1994), have expressed this system similarly as $y_j = k$ if $\tau_{j,k-1} < y^*_j \leq \tau_{j,k}$, where the thresholds values for the $j^{th}$ item are: $-\infty = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_{k-1} < \tau_k = \infty$. The lower ($\tau_0$) and upper ($\tau_k$) are bounds and not needed practically. Both manners of expressing the categorization of the latent response continuum are equivalent. Each item can have possibly unique threshold(s) that are estimated. However, as the number of categories increases so does the computational burden of estimating thresholds, meaning that the number of parameters estimated increases systematically as the number of items and number of categories increases. Note that for continuous indicators, the above system is unnecessary as the observed score is equal to the underlying latent response score, i.e. $y_{ij} = y^*_{ij}$. This formulation of categorical data is sometimes called the latent variable formulation (L. Muthén & Muthén, 2017). Ordered categorical data are known to causes issues in the estimation of factor models. Below is a brief review of the literature on the effect of categorical data on CFA estimation.

*Parameter Estimates*

In the estimation of model parameter, the covariances among items must be estimated, and when coarsely defined variables are used, the covariance is attentuated (Bollen & Barb, 1981; Johnson & Creech, 1983). These finding have been consistently replicated across many scenarios through simulations studies when the known correlation among items and the effect of categorization can be explicitly examined. Overall, when the Pearson Product-Moment correlation coefficient is used to approximate the relationships among items the parameter estimates tend to be negatively biased (Babakus, Ferguson, & Joreskog, 1987; Muthén & Kaplan, 1985). However, the impact on the effect depends on the choice of estimator, e.g. maximum likelihood, weighted least squares or diagonally weighted least squares (Bandalos, 2014; Muthén & Kaplan, 1985; Muthen & Kaplan, 1992). The amount of bias is also relatively small under most conditions, but the bias increases in magnitude as the severity of nonnormality increases (Bandalos, 2014; DiStefano & Morgan, 2014; Forero, Maydeu-Olivares, & Gallardo-Pujol, 2009; Li, 2016).

*Standard Errors*

The literature is not consistent about the effect of categorical data on standard errors (SE). Through simulation studies, some authors have found a consistent overall negative bias (Bandalos, 2014; DiStefano & Morgan, 2014; Dolan, 1994; Muthén & Kaplan, 1985). Others have identified a positive bias (Babakus et al., 1987), and others still have identified a mix of bias that depends on the severity of nonnormality (Forero et al., 2009; Li, 2016; Nestler, 2013).

*Model Fit*

One of the major issues that arise when using CFA is establishing whether the measurement model sufficiently explains the relationship among observed variables. That is the hypothesized model is expected to completely explain the intercorrelations among items. However, the measurement model aims to be as parsimonious as possible. In educational setting, the aim is often to measure a single construct with a set of items (e.g., math ability). The generally accepted guidelines for establishing model fit for CFA models is based on fit statistics such as the $\chi^2$ test of model fit, confirmatory fit index (CFI), Tucker-Lewis Index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). The model $\chi^2$ is known to be sensitive to sample size (Bollen, 1989). However, some authors have found with robust estimators that the model $\chi^2$ to be close to expected values (Bandalos, 2014; Flora & Curran, 2004). The other fit statistics have the recommended cut-off values of CFI $>$ .95, TLI $>$ .95, RMSEA $<$ .06, and SRMR $<$ .06 (Hu & Bentler, 1999). The Hu and Bentler (1999) cut-off values have been called into question over the past two decades as being established under too limited of conditions and not with categorical data (Finney & DiStefano, 2013). Some authors have even recommended that rules of thumb for robust estimators for categorical data cannot be established (Nye & Drasgow, 2011, as cited in Finney, 2013). In general though, the interpretation of common fit statistics from robust estimators has not been well established.

*Recommended Estimators*

For ordered polytomous response data, the following estimators are suggested for different scenarios. The general recommendation is to treat the observed data as categorical and use a robust DWLS estimator (DiStefano & Morgan, 2014; Finney & DiStefano, 2013). When the number of categories is five or more, there is evidence that treating the data as continuous and using a robust estimator is sufficient for obtaining unbiased estimates of parameters in some circumstances (Finney & DiStefano, 2013; Rhemtulla, Brosseau-Liard, & Savalei, 2012). Other authors also found evidence that using ULS can provide more accurate parameter and SE estimates when the estimation converges (Forero et al., 2009). Due to the large number of available estimators for authors to choose from, Finney and DiStefano (2013, pg. 476) also recommended fitting the model with more than one estimator to check for converging evidence of one's solution.

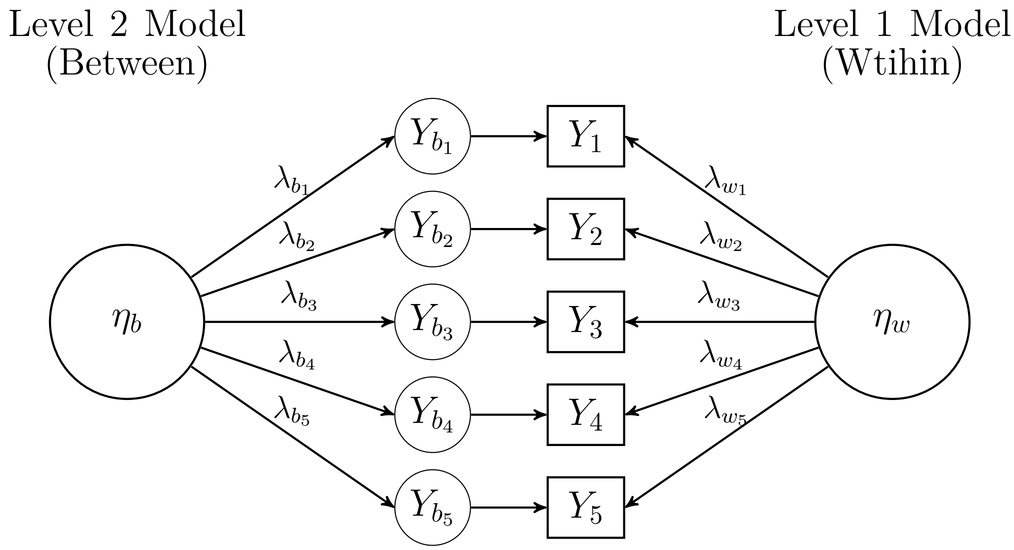
*Multilevel Confirmatory Factor Analysis*

Similar to how methodological work has expanded to address the categorical nature of item responses in CFA, multilevel modeling has also been incorporated into CFA. One of the assumptions underlying traditional single level CFA is that the individuals being sampled are independent of one another. However, as explored earlier in this chapter, educational data rarely meet this assumption given that data often arise from nested structures. Two general approaches can be taken to account for the dependence among cases. The first is based on a sampling theory perspective. The second is a varying parameter modeling perspective. In the first framework, traditional single level analysis are conducted, but the standard errors are corrected for the sampling design. For example, one could take a random sample of schools and then within each school take a random sample. Because we know we are taking random sample within sampling units, we can account for the sampling design based on adequately developing sample weights; this is why this approach is sometimes referred to as a "design based" approach (Kaplan, 2009; Stapleton, 2013; Stapleton, Yang, & Hancock, 2016). In the second approach, parameters are treated as random components that vary across higher level units. When using this approach, the researcher explicitly models the mechanism by which parameters vary across schools. Utilizing this approach accounts for the dependence among cases within unit by modeling and estimating the differences among schools. This latter approach is known as multilevel confirmatory factor analysis (MCFA) and is the focus of this work.

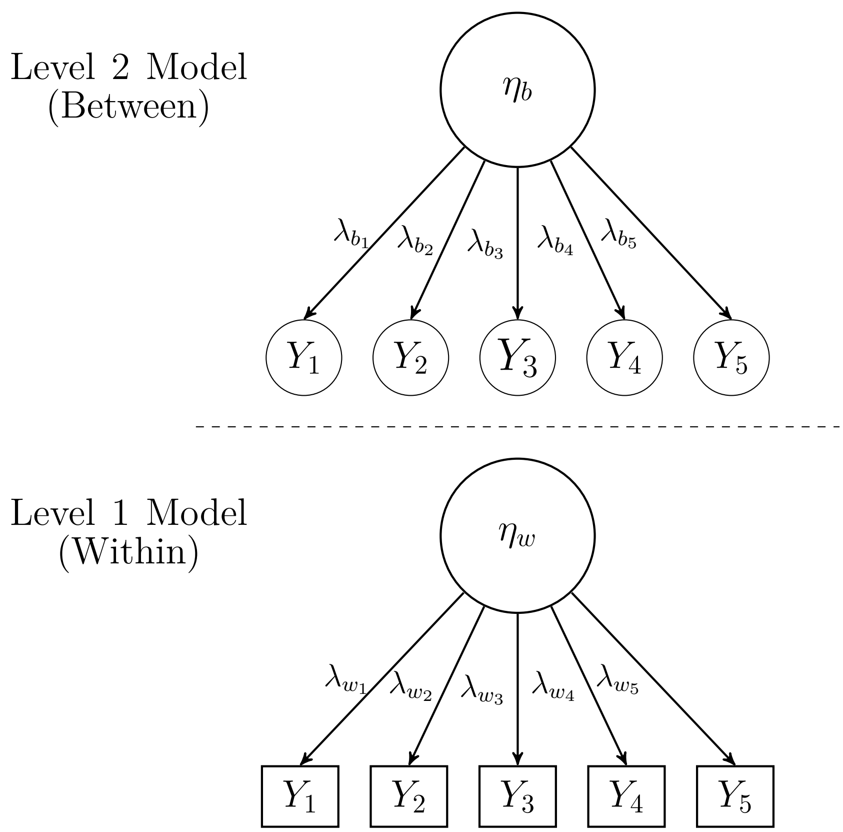
In his seminal work on extending the conventional CFA model for multilevel settings, Muthén (1994) showed how the observed covariance matrix can be parceled into a pooled-within cluster covariance matrix (i.e. level-1 covariances) and a between cluster (i.e. level-2 covariances) covariance matrix to model the nested structure of data. The pooled-within cluster covariance matrix defines how the item responses covary because of the individual contributions of each respondent. For the remainder of this paper, the pooled-within cluster covariance matrix are referred to as within covariances or level-1 covariances. In contrast, the between cluster covariance matrix defines how the average response to each item covaries across groups. For the remainder of this paper, between cluster covariance matrix are referred to as between covariances or level-2 covariances. During the early applications of these models, researchers computed the level specific covariance matrices and estimated the model in a multiple framework. B. Muthén expanded upon the work of Goldstein and McDonald (1988), McDonald and Goldstein (1989), and B. Muthén and Satorra (1989) by showing how to efficiently estimate these complex models in one framework.

In MCFA, the covariance structure is decomposed into the level-1 (within) and level-2 (between) covariance matrices with corresponding factor structure specification. The level-1 model represents the individual/student level processes that influence the observed responses. The level-2 model represents the factors influencing variations in responses across groups or schools. One of the difficulties in utilizing MCFAs is establishing the meaning of the level-2 constructs because the meaning depends on how the factors relate to the level-1 factors. A full description of the technical issues surrounding construct meaning for level-2 factors is out of scope for the paper, and the interested reader is referred to the excellent article by Stapleton et al. (2016).

The MCFA model is expressed as in a path model at each level of analysis, namely level-1 and level-2. An example of a simple two level MCFA with five indicators is shown in Figure 2.2. The reader will notice that the model structure is equivalent across both levels. For example, Muthén (1994) representation explicitly shows the effect of the level-2 model on the observed indicators (see Figure 2.2). When the MCFA is specified with equal structure and factor loadings are constrained equal across levels, the researcher can appropriately calculate the proportion of variation explained in the latent variable ($\eta$) by group membership can be estimated (i.e., the intraclass correlation). However, a precursor to estimating the ICC in the latent




*(a)* Muthén (1994) Path Model




*(b)* Alternative MCFA Path Model

*Figure 2.2.* Path Models for Multilevel Confirmatory Factor Analysis

variables is establishing fit of the measurement model across both levels (Heck & Thomas, 2015; Stapleton, 2013). Due to the complexity of MCFA models, researchers have suggested a four-step[1] approach to evaluating one's model (Dyer, Hanges, & Hall, 2005; Muthén, 1994; Stapleton, 2013).

(1) Fit a single level CFA based on a priori expectations of model structure

(2) Estimate the corrected level-1 covariance matrix and estimate the level-1 model

(3) Estimate the corrected level-2 covariance matrix and estimated the level-2 model

(4) Estimate the full model with level-1 and level-2 specified based on theoretical expectations

These general four steps are designed to help the researcher identify potential model specification issues at each level. The technical details on the estimation of the level specific covariance matrices can be found in Kaplan (2009); Muthén (1994). A relatively simple introduction to these models can be found in Dyer et al. (2005); Kaplan (2000); Kline (2016), where each author gives an example application. In what follows, I describe the limited methodological literature that tests the limits of the MCFA.

*Paramater Estimation*

The majority of the methodological work on MCFA investigated conditions in which the factor indicators are normally distributed (Hox & Maas, 2001; Wu & Kwok, 2012). Precision of estimation of model parameters in MCFA depends on number of level-2 units (e.g. number of schools), average number of observations within each level-2 unit, distribution of indicators, and estimator (Hox & Maas, 2001). The level-1 parameters have been found to be relatively unbiased under most conditions when

[1] Sometimes referred to as a five step approach where the estimation of the two covariances matrices are step 2. Steps 3 and 4 are estimating the level-1 and level-2 models, and step 5 is estimating the full model.

data are multivariate normally distributed regardless of sample size (Asparouhov & Muthén, 2007; Hox & Maas, 2001). However, parameter bias increased rapidly for level-2 specific parameters, where all parameters, especially factor loadings, showed stronger downward bias (Asparouhov & Muthén, 2007; Hox & Maas, 2001; Wu & Kwok, 2012). To overcome the bias, authors have suggested at least 100 level-2 units (Hox & Maas, 2001).

*Standard Errors*

At the within part of the model, SE did not have a reported relative bias that is concerning across any conditions, only 0.6% as most (Hox & Maas, 2001). However, at level-2, SE showed a clear downward bias for factor loadings across all conditions (-9.0% on average for even the largest Ng condition of 200). The downward bias effect was reduced for lower levels of ICCs. Hox & Maas do did not find the level-2 variances to be significantly biased in either direction.

Note that parameter coverage rates were below the nominal level for factor loadings and variances and the coverage was worse for unbalanced group sizes.

*Model Fit*

A very limited literature exists on the performance of fit indices in MCFA with categorical data (Hsu, 2009; Navruz, 2016). However, a few additional studies have investigated fit indices in MCFA when indicators are multivariate normally distributed (Asparouhov & Muthén, 2007; Hox & Maas, 2001; Hsu, 2009; Hsu, Kwok, Lin, & Acosta, 2015; Ryu & West, 2009; Wu & Kwok, 2012). Under data simulated as multivariate normally distributed, the consensus so far is that CFI, TLI, RSMEA, SRMRW are only sensitive to misspeficiation of the level-1 model, while the SRMRB is able to detect misspecification of the level-2 model. However, the study that has most thoroughly investigated fit in MCFA is limited in generality due to the methods employed (Hsu et al., 2015). Hsu et al. (2015) found that the Hu and Bentler (1999) cut-off criteria for CFI, TLI, and RSMEA performed well for detecting misspecification at level-1 but not when only level-2 is specified. However, when using categorical data with the WLSMV or WLSM estimator, Navruz (2016) found that none of the fit statistics were able to consistently identify the correct model with the traditional cut-offs.

*Model Estimation*

Of the four studies that have investigated multilevel CFA with categorical indicators, authors have invested three estimators: MLR, WLSM, WLSMV. Asparouhov and Muthén (2007) recommended using WLSM over MLR for estimation due to WLSM resulting in less biased estimates across conditions studied. However, only one condition was investigated. While in Study 2 of Hsu (2009), only WLSMV was used to investigate the effects of dichotomous indicators on MCFA. In a small simulation study comparing cluster-effects on parameter recovery, Stochl et al. (2016) used the WLSMV estimator in a condition where a MCFA model was estimated with WLSMV to compare results of level-1 parameters to design-based approaches. In Navruz (2016), the author compared WLSM with WLSMV and concluded that, under the conditions included, neither estimator should be used for ordered categorical data in multilevel settings.

## *Purpose and Research Questions*

Limited research exists that examine multilevel structural equation models, and especially under conditions when data are ordered categorical. Given that data from educational settings typically arise from hierarchically structured contexts and used ordered categorical response formats, there is a need to understand how these models can be examined. Establishing that one's measurement model adequately fits the observed data is therefore a central issue that needs to be considered in these complex scenarios. Due to the limitations of prior studies of the available fit indices for MCFA models, one of the aims of this paper is to investigate whether a different cut-off value would be able to detect a global misfit of the level-1 or level-2 model. The purpose of this thesis is therefore to examine how fit statistics perform in detecting misspecified measurement models in multilevel CFAs. Specifically, my research questions are

(1) Is there a cut-off for CFI, TLI, RSMEA, SRMRW, and SRMRB that can be used to differentiate a properly specified model from an misspecified model?

(2) If there is a cut-off, what is the values of each of these fit indices?

(3) Finally, how does identification of fit vary across potential estimators for MCFA models?

# CHAPTER THREE

## Methods

In this chapter, the Monte Carlo simulation methods employed in this study are described. A cursory introduction to the technical details of MCFA models and the estimation of these models are given. Additionally, the final simulation study outcomes being analyzed are explained in detail along with the final statistical analyses that are reported.

### *MCFA Model*

As briefly described earlier, the multilevel confirmatory factor analysis (MCFA) model is a decomposition of the observed covariance matrix into a level-1 (pooled within group) and level-2 (between group) specific covariance matrix. Because two covariance matrices are essentially being analyzed, two possibly unique models can be specified for each model. These two models can be broken down to be two components that influence the each individual response. First, the individual effect and a second group effect. With categorical observed data, these two effects directly influence the underlying response value for each item, as shown in Equation 3.1. In this equation the underlying response value for an individual ($y^*_{pig}$) is the composition of the individual component ($y_{wpij}$) and a random effect of group ($y_{bpj}$). As with Muthén (1984) and Asparouhov and Muthén (2007), the latent response is assumed normally distributed.

$$y^*_{pig} = y_{bpj} + y_{wpij} \tag{3.1}$$

where $y^*_{pig}$ is the latent response underlying the observed categorical value $y_p$ for the $p^{th}$ item, $i$ indexes across individuals within group $g$. Across $p$ items, this general framework is easily incorporated into the previous notation for factor analysis by replacing the latent components $y_{wpij}$ and $y_{bpj}$ by the level specific factor loading matrices and latent variable vectors. This can be seen more explicitly by looking at the factor analytic function being estimated in Equation 3.2.

$$\mathbf{y}^*_{\mathbf{ig}} = \nu_g + \mathbf{\Lambda}_B \eta_{Bg} + \varepsilon_{Bg} + \mathbf{\Lambda}_W \eta_{Wig} + \varepsilon_{Wig} \tag{3.2}$$

where $\mathbf{y}^*_{\mathbf{ig}}$ is the latent response vector for the $i^{th}$ individual in the $g^{th}$ group, $\nu_g$ are the group latent intercepts (which are not directly estimated), $\mathbf{\Lambda}_B$ is the matrix of factor loadings for the level-2 (between) model, $\eta_{Bg}$ is the vector of group latent variable scores, $\varepsilon_{Bg}$ is the vector of group level residuals, $\mathbf{\Lambda}_W$ is the matrix of factor loadings for the level-1 (within) model, $\eta_{Wig}$ is the vector of individual latent variable scores, and $\varepsilon_{Wig}$ is the vector of individual level residuals for individual $i$ in group $g$. Note that the observed scores are categorical in this case, so item threshold are used to relate the underlying response values ($\mathbf{y}^*_{\mathbf{ig}}$) to the observed vector of categorical responses ($\mathbf{y}_{ig}$). These thresholds ($\tau_{pk}$) are specific to each item (p) and constant across groups. Meaning the the relationship described in Equation 2.5 is simply expanded to include the categorization across groups (see Equation 3.3). The total number of categories in the observed variable is defined as k.

$$y_{pig} = k \Leftrightarrow \tau_{pk} < y^*_{pig} < \tau_{p(k+1)} \tag{3.3}$$

where $\tau_{p0} = -\infty$ and $\tau_{pk+1} = \infty$. The relationship between the observed categories and underlying response vector allows for the categorical nature of the observed data to be controlled for when estimating the between group variability.

The total variation ($\mathbf{\Sigma}_T$) is then expressed as a summation of the within and between group variability.

$$\begin{aligned} \mathbf{\Sigma}_T &= \mathbf{\Sigma}_B + \mathbf{\Sigma}_W \\ &= \mathbf{\Lambda}_B \mathbf{\Psi}_B \mathbf{\Lambda}_B^{\mathrm{T}} + \mathbf{\Theta}_B + \mathbf{\Lambda}_W \mathbf{\Psi}_W \mathbf{\Lambda}_W^{\mathrm{T}} + \mathbf{\Theta}_W \end{aligned} \tag{3.4}$$

where $\mathbf{\Sigma}_W$ is the within group covariance matrix, $\mathbf{\Sigma}_B$ is the covariance among group means, $\mathbf{\Psi}_W$ is the covariance matrix among level-1 latent variables, $\mathbf{\Psi}_B$ is the covariance among level-2 latent variables, $\mathbf{\Theta}_W$ is the level-1 residual covariance matrix, and $\mathbf{\Theta}_B$ is the level-2 residual covariance matrix among group means. The level-1 and level-2 covariance matrices are assumed to be independent, meaning that how much error is associated with measurement of group means is not related to the amount of error in measuring item responses. In other words, the sampling process does not induce a dependence among school means and responses by students. In general, Equation 3.4 implies that the covariance among the observed indicators is decomposed between the within measurement model and the between groups measurement model. By specifying the covariation at both levels of sampling, we gain the flexibility to describe relationships at the individual level and the group level. The measurement of traits across levels of analysis allows researchers to test detailed theories of organizational effects and measurement concerns at each level of sampling.

*Estimation*

The estimation MCFA models is determined in part of the estimation of two covariance matrices simultaneously for the corrected level-1 (within, $\Sigma_W$) and the corrected level-2 (between, $\Sigma_B$) covariances. The population covariance matrices cannot be directly estimated. Muthén (1989, 1990) showed how the following sample covariance matrices can be used together to estimated the population covariances.

$$\mathbf{S_W} = (N-G)^{-1}\sum_{g=1}^{G}\sum_{i=1}^{N_g}\left(\mathbf{y_{ig}}-\bar{\mathbf{y}}_\mathbf{g}\right)\left(\mathbf{y_{ig}}-\bar{\mathbf{y}}_\mathbf{g}\right)' \tag{3.5}$$

$$\mathbf{S_B} = (G-1)^{-1}\sum_{g=1}^{G}N_g\left(\bar{\mathbf{y}}_\mathbf{g}-\bar{\mathbf{y}}\right)\left(\bar{\mathbf{y}}_\mathbf{g}-\bar{\mathbf{y}}\right)' \tag{3.6}$$

where $N$ is the total number of observations, $G$ is the total number of groups or level-2 units, $N_g$ is number of individuals in group $G$, $\mathbf{y_{ig}}$ is the observed vector of response of individual $i$ in group $g$, $\bar{\mathbf{y}}_\mathbf{g}$ is the vector of average responses for group $g$, and $\bar{\mathbf{y}}$ is the vector of average responses for the whole sample. The between (level-2) covariance matrix is described as the sample-size weighted covariance among group means (Muthén, 1994). For more nuanced technical details on the estimation of each of these matrices and the fit function being optimized, the reader is referred to the

M*plus* technical web note (B. Muthén, du Toit, & Spisic, 1997) and the LISREL technical appendix (Jöreskog, 2005).

*Model Estimators*

All models were estimated using M*plus* (L. Muthén & Muthén, 2017, version 8.2). Researchers have the choice of a many potential estimators in M*plus*, and each was initially developed to solve a particular estimation limitation of another estimator. A full list of the available estimators for models in M*plus* can be found in the User's Guide (L. Muthén & Muthén, 2017). In this project, I focus on ordered categorical responses, for which the maximum likelihood with robust standard errors (MLR) and weight least squares mean and variance adjusted (WLSMV) estimators have been tested for MCFA models. However, for single level CFAs with categorical data, there is evidence that unweighted least squares (ULS) performs well (Forero et al., 2009). But, the ULS estimator is not available for multilevel models in M*plus*, so instead, the unweighted least squares mean and variance adjusted (ULSMV) was used. Therefore, I focus on three estimators: MLR, WLSMV, and ULSMV. Each of these estimators aims to optimize a fit function that describes the discrepancy between the observed sample (co)variances and the model implied (co)variances.

*MLR.* The maximum likelihood with robust standard error (MLR) estimator is a two part estimation process. In the first part, traditional maximum likelihood is performed to obtain parameter estimates. The second part uses a sandwich estimator to estimator standard errors. The interested reader is referred to Bandalos (2014) for a more in-depth discussion of MLR with categorical indicators. With categorical data, MLR requires the use of numerical integration for each latent variable in the model. This feature makes MLR computational intensive for more than four latent factors. In MCFA, most models of interest in practice would require at more than four latent variables. Therefore, when using the MLR estimator, I am setting the observed indicators to be read as continuous measures. Ordered response formats with at least five categories does have evidence for being treated as continuous in single level analyses (Rhemtulla et al., 2012).

*ULSMV.* Unweighted least squares (ULS) estimator uses the identity matrix to weight the least-squares portion of the optimization function. This means that each (co)variance component in the model contributes equally base on magnitude of discrepancy to the fit of the solution In M*plus*, the ULS estimator is only available for the single-level analyses. However, for two-level analysis this estimator is not available, but an alternate form of ULS is available: unweighted least-squares parameter and standard error estimates with mean and variance adjusted $\chi^2$ statistic (ULSMV).

*WLSMV.* The weight least squares mean and variance adjusted (WLSMV) estimator is a form of diagonally weighted least-squares (DWLS). DWLS estimators optimize the discrepancy function while weighting each (co)variance component discrepancy based on the covariance among the (co)variance components. The DWLS estimator was designed to account for the categorical nature of ordinal responses in order to more precisely estimate the (co)variance components among ordered categorical responses.

### *Model Fit Indices*

For factor analysis models, researchers find evidence for whether the hypothesized relationships among items are representative of the observed data through fit indices. Fit indices are broadly measuring how well a model reproduces the observed relationships or how poorly the relationships are capture by the model, i.e. a measure of goodness of fit versus a measure of badness of fit. If a researcher's hypothesized measurement model is able to represent and recapture the inter-correlations among item, then the researcher has evidence that the hypothesized measurement model is appropriately specified. Below each of the measures of fit of interest in my project that are innately reported by M*plus* v8.2 when estimating MCFA models.

*Measures of Goodness of Fit*

The measures of goodness of fit are indicative of how much better of fit the hypothesized model provides over a null model. A null model refers to a measurement model where all items are assumed independent, which is the worst-case scenario and generally is the worst fitting model. The first measure of goodness of fit is the $\chi^2$ statistic and the associated $\chi^2$ test. Each model estimated has an associated model $\chi^2$, which is a statistical distribution that has an expected value equal to model degrees of freedom. The $\chi^2$ test is known to be sensitive to sample size, meaning that as sample size increases the test is likely to reject the null hypothesis that these data fit the hypothesized model even when the model is correctly specified (Bollen, 1989). Because the $\chi^2$ statistic has limited applicability, numerous other statistics based on the $\chi^2$ are frequently used. These statistics are transformations of the $\chi^2$ statistic that have seen broad applicability.

*CFI.* The comparative fit index (CFI) is a commonly used fit statistic that is based on the model $\chi^2$ (Bentler, 1990). The CFI is a measure of improvement in fit over the null model with a fixed range of zero to one, where higher scores mean better fit.

$$CFI = 1 - \frac{\max\left(\chi^2_H - df_H,\ 0\right)}{\max\left(\chi^2_H - df_H,\ \chi^2_N - df_N,\ 0\right)} \tag{3.7}$$

The recommended minimum value for CFI is .95 (Hu & Bentler, 1999).

*TLI.* The Tucker-Lewis index (TLI) is another measure of fit over the null model that is non-normed (Bentler & Bonnett, 1980; Tucker & Lewis, 1973). The TLI is non-normed because the range of possible values is not restricted, where values greater than (or less than) one are possible. Although, TLI values above one are typically rescaled to one in practice. But, just like CFI, higher values of TLI are indicative of better fit.

$$TLI = \frac{\frac{\chi^2_N}{df_N} - \frac{\chi^2_H}{df_H}}{\frac{\chi^2_N}{df_N} - 1} \tag{3.8}$$

The recommended minimum value for TLI is .95 (Hu & Bentler, 1999).

*Measures of Badness of Fit*

Of interest in this project, are a few measures of lack of fit or how badly the hypothesized model fits the observed data. Three measures are of interest: the RMSEA, SRMRW, and SRMRB. Smaller values on these fit statistics are indicative of better fit.

*RMSEA.* The root mean square error of approximation (RMSEA) is a measure of lack of fit (Steiger & Lind, 1980). The RMSEA describes how well the hypothesized model is represented in the observed data. The comparison is accomplished by a measure of the deviation of what is observed

$$RMSEA = \sqrt{\frac{\max\left(\chi^2 - df,\ 0\right)}{df(N-1)}} \tag{3.9}$$

The RMSEA has a known population distribution. Because the distribution is known based on a noncentral F-distribution, a confidence interval can be constructed. Usually, a 90% confidence is reported for the RMSEA. The recommended minimum value for RMSEA is .06 (Hu & Bentler, 1999).

*SRMR.* The standardized root mean square residual (SRMR) is an aggregate measure of the deviation of the observed correlation matrix to the model implied correlation matrix (Jöreskog & Sörbom, 1981). Ideally, the average difference between the observed and expected correlations is minimal, and smaller values represent better fitting models. In MCFA, two difference SRMR are reported based on each level's covariance matrix. The level-1 SRMR is known as SRMR-Within (SRMRW). The level-2 SRMR is known as SRMR-Between (SRMRB). The computation of each of these measures is roughly equivalent conditional on which covariance matrix is under consideration. Generally, SRMR is computed as the standardized difference between the observed correlations and the model implied correlations about variables as shown below.

$$SRMR = \sqrt{\frac{2\sum_{j=1}^{p}\sum_{k=j}^{i}\left(\frac{s_{jk}-\hat{\sigma}_{jk}}{\sqrt{s_{jj}s_{kk}}}\right)^2}{p(p+1)}} \tag{3.10}$$

where $p$ is the total number of variables in the model, $s_{jk}$ and $\sigma_{jk}$ are the sample and model implied, respectively, covariance between the $j^{th}$ and $k^{th}$ variables. For the SRMR, generally acceptable values less than .08 are used. However, Hu and Bentler (1999) suggested values less .06 alone or .08 in combination of other within recommended ranges are indicative of good fit.

In M*plus*, SRMR is estimated slightly differently than shown in Equation 3.10. The computation extends the definition above by accounting for the meanstructure, multilevel structure, and categorical nature if applicable. The technical details of the computation of SRMRW and SRMRB in M*plus* v8.2 is out of scope for this article, and the interested reader is referred to B. O. Muthén and Asparouhov (n.d.) for more details.

### *Data Generating Process*

Data for this project were generated using M*plus* v.8.2 (L. Muthén & Muthén, 2017) utilizing the MONTECARLO command along with MplusAutomation (Hallquist & Wiley, 2018) package in R (R Core Team, 2018).

### *Fixed Design Factors*

Data were generated from a factor model with 10 items and two correlated factors at both levels. Across levels, the structure and factor loadings were specified to be invariant. The factor loadings were set to 0.60, which corresponds to the lower end of strong factor loadings (DiStefano & Hess, 2005). At level-1, factors were generated with a fixed unit variance. Fixing the factor variance to one made specifying the level-1 correlation simple, meaning all that is needed is specifying the covariance to the correlation. The factors at each level were specified to have a correlation of .3, similar to Hsu et al. (2015). The level-2 factor (co)variances depended on the ICC condition, and more detail is given in the ICC section of manipulated factors. The correlation is constant, but the specific value of (co)variance at level-2 changed across conditions.

Figure 3.1 shows the population structure of the MCFA simulated for one of the ICC conditions. The two factors were measured with 6 items and 5 items respectively. Notice that one indicator is cross-loaded on both factors at both levels in the population. Previous studies of model fit measures have typically generated data from a population structure with at least one cross-loading so that the effects of misspecification can be examined, for examples see Hsu (2009); Hsu et al. (2015); Hu and Bentler (1998, 1999); Navruz (2016); Wu and Kwok (2012).

*Indicator Distribution*

The observed indicators were generated to be ordered categorical, e.g. Likert-type responses. One of the difficulties in working with categorical indicators in MCFA is correctly specifying the variance of the observed indicators. For factor models, there is an interdependence between the factor variance, factor loading(s), indicator residual variance, and total variance of of the latent response continuum. The relationship between a single factor and a single underlying continuum in a single level factor analysis is expressed in Equation 3.11

$$Var(y_i^*) = \lambda_i^2\phi + \psi_{ii} \tag{3.11}$$

where $Var(y_i)$ is the total variance of item $i$, $\lambda_i^2$ is the factor loading of item $i$ with factor $\phi$, and $\psi_{ii}$ is the residual variance of item $i$. The scale of latent response

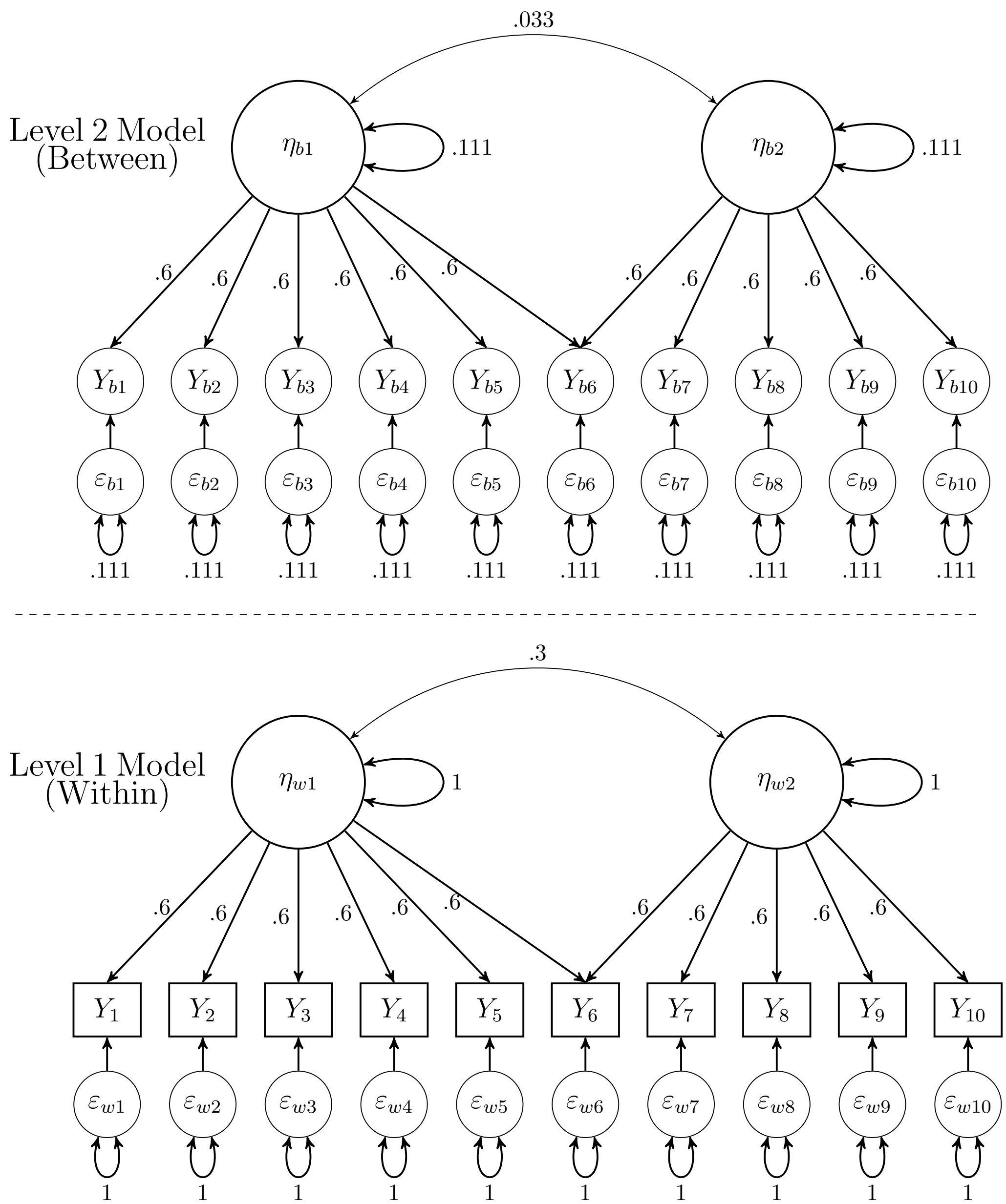



*Figure 3.1.* Simulation data generating model specification.
*Note.* The observed variance of each indicator is fixed to one across all items with an observed and latent ICC of .1.

continuum can be set by either fixing the total variance of the fixing the residual variance.

This relationship extends to MCFA, and is used to set the observed variance based on factor loadings and residual variances. The total variance of $y^*$ for a single item with no cross-loadings in MCFA was found with

$$Var(y^*) = \lambda_w^2 \phi_w + \psi_w + \lambda_b^2 \phi_b + \psi_b \tag{3.12}$$

where each of the components above are analogous to the single level pieces but for the level-1 and level-2 models. This information was used to establish the observed variance of each item.

The observed variance was controlled by adjusting the thresholds for the categories. The items were generated with approximately unit variance and bell-shape distribution based on the thresholds for latent response variables. I chose this distribution to represent ideal conditions of polytomous response data given the limited research in this area. Variances for discrete response sets can be found in the population using the following method.

(1) Calculate the proportion of responses, i.e. response probability, for each ordered category

(2) Estimate the squared expected value (squared average, $[E(X)]^2$ ) by multiplying the probability of each category by the coded value for that category, i.e. 0, 1, 2, etc., and then sum across categories. Then square this sum.

(3) Estimate the expected value of the squared scores ($E(X^2)$) by multiplying the probability of each category by the coded value for that category squared, i.e. $0^2$, $1^2$, $2^2$, etc., and then sum across categories.

(4) The difference between the expected value of the squared scores and the squared expected value is the variance ($Var(X) = E(X^2) - [E(X)]^2$).

The above steps allowed the variance to be estimated based on the probability of responding to each category. The response probabilities are shown in Table 3.1. Based

Table 3.1

*Response probabilities for simulating unit variance*

| Response Category | Value | Probability | Threshold[1] |
|---|---|---|---|
| Strongly Disagree | 1 | 0.0675 | |
| Disagree | 2 | 0.2325 | -1.95 |
| Neutral | 3 | 0.4000 | -0.68 |
| Agree | 4 | 0.2325 | 0.68 |
| Strongly Agree | 5 | 0.0675 | 1.95 |

*Note.* [1] Thresholds for item 1 with ICC of .2, and variance of $y^*$ is 1.7.

on the response probabilities, the threshold values necessary for categorizing the latent response continuum were computed. The latent response continuum is assumed normally distributed with a mean of zero and standard deviation $(N(0, \sqrt{Var(y^*)})$ prior to being categorized. The $Var(y^*)$ is found based on the simulation condition of level-2 variances. An R script was written to compute the thresholds based on the design factors and is given in Appendix A. Because five categories were simulated, four thresholds are needed to categorize the latent response continuum. The four threshold categorize the continuum based on the scheme given in Equation 2.5. Unique thresholds were needed for item 6 because of the cross-loading and for different conditions because of varying levels of residual variance at level-2. The thresholds for each ICC condition are reported in Table 3.2.

*Manipulated Design Factors*

Factors that varied across simulation conditions are indicator ICC, latent variable ICC, number of groups, and number of level-1 units.

Table 3.2

*Model generating thresholds across all conditions*

| Condition | | | | Thresholds | | | |
|---|---|---|---|---|---|---|---|
| $ICC_L$ | $ICC_O$ | Item | Variance[1] | $\tau_1$ | $\tau_2$ | $\tau_3$ | $\tau_4$ |
| .1 | .1 | 1 | 1.509 | -1.83644 | -0.64431 | 0.64431 | 1.83644 |
| .1 | .1 | 6 | 2.148 | -2.19109 | -0.76874 | 0.76874 | 2.19109 |
| .1 | .3 | 1 | 1.829 | -2.02173 | -0.70932 | 0.70932 | 2.02173 |
| .1 | .3 | 6 | 2.468 | -2.34857 | -0.82399 | 0.82399 | 2.34857 |
| .1 | .5 | 1 | 2.399 | -2.31534 | -0.81233 | 0.81233 | 2.31534 |
| .1 | .5 | 6 | 3.038 | -2.60560 | -0.91417 | 0.91417 | 2.60560 |
| .5 | .1 | 1 | 1.830 | -2.02196 | -0.70940 | 0.70940 | 2.02196 |
| .5 | .1 | 6 | 2.982 | -2.58107 | -0.90556 | 0.90556 | 2.58107 |
| .5 | .3 | 1 | 2.150 | -2.19162 | -0.76892 | 0.76892 | 2.19162 |
| .5 | .3 | 6 | 3.302 | -2.71603 | -0.95291 | 0.95291 | 2.71603 |
| .5 | .5 | 1 | 2.720 | -2.46508 | -0.86486 | 0.86486 | 2.46508 |
| .5 | .5 | 6 | 3.872 | -2.94113 | -1.03188 | 1.03188 | 2.94113 |

*Note.* [1] Variance of latent response variable prior to categorization. $ICC_L$ is the ICC for the latent variables, $ICC_O$ is the ICC for the observed variables.

*Intraclass-Correlation*

The intraclass-correlation (ICC) is the proportion of variance accounted for by group membership; in other words, how much variability in the observed scores is due differences in how groups response to the questions on average. The ICCs were manipulated by fixing the residual variance at level-1 to one and changing the level-2 variance, which is a similar approach to Navruz (2016). In M*plus* this was accomplished by using the command: PARAMETIZATION = THETA. The default parameterization for ordered categorical data is DELTA, which fixes the latent response distribution to have unit variance instead of the residual variance. Hsu (2009) established the ICCs using the DELTA parameterization when simulating dichotomous indicators, and manipulated the values of the ICCs using the factor loadings. The ICC for the is calculated as proportion of level-2 variance to level-1 variance

(Heck & Thomas, 2015, pg. 159).

$$\text{ICC} = \frac{\psi_b}{\psi_b + \psi_w} \tag{3.13}$$

where $\psi_b$ is the variance for level-2 and $\psi_w$ is the variance for level-1.

*Observed Variable ICC.* In this study, three different ICC values .1, .3, and .5 were investigated for each indicator ($\text{ICC}_O$). Each of the indicator ICCs were manipulated by changing the residual variance at level-2 of each item. The level-2 item variances were .11, .43, and 1 for each of the ICC conditions, respectively.

*Latent Variable ICC.* Furthermore, the latent variable ICCs ($\text{ICC}_L$) were manipulated. Values of .1 and .5 were investigated corresponding to a low and high level.

### *Number of Groups*

The number of groups ($\text{N}_2$) is known to be one of the most influential components to estimation of MCFA. For this investigation, four different group sizes is under investigation, namely 30, 50, 100, and 200.

### *Number of Level-1 Units*

The number of units sampled within each group ($\text{N}_1$) was held constant for simplicity with three different levels. The number of units varied among 5, 10, and 30.

## *Summary of Model Fitted*

The major design aspect of this study is the identification of the correctly specified model amongst misspecified models. For each generated dataset, four different model specifications were fit, namely

(1) Level-1 and level-2 measurement model are correctly specified (C)

(2) Level-1 measurement model is misspecified (M1)

(3) Level-2 measurement model is misspecified (M2)

(4) Level-1 and level-2 measurement models are misspecified (M12)

The misspecification in each of these levels omits the cross-loadings at the desired level(s).

In additional to model specification, three different estimators were tested. This creates an additional layer of models that were estimated within each condition. The MLR, ULSMV, and WLSMV estimators were each used on the datasets generated in each condition.

*Influence of Design Factors*

The influence of manipulated design factors on the distributions of each fit statistic (i.e. CFI, TLI, RMSEA, SRMRW, and SRMRB) was assessed with two-way factorial ANOVAs. Using an experimental design framework to summarize the results of Monte Carlo simulation studies provides a rigorous framework for describing the large quantities of data (Harwell, 1992). For each fit statistic, the overall influence of all interactions and main effects were assessed with $\omega^2$ and partial-$\omega^2$ (Maxwell & Delaney, 2004). This analysis design is extended for each estimator.

*ROC Analysis for Fit Index Cut-Off*

The cut-off values established by Hu and Bentler (1999) were determined, in part, by an inspection of the rejection rates of various cut-off criteria. In some conditions, the rejection rates were based on a rule with two different fit statistics. A rejection rate is the proportion of observed fit statistic value was outside of the cut-off criteria range. This value is also sometimes called a hit-rate, and the inverse is called the acceptance rate. In this study, I investigated acceptance rates for correctly specified models across a range of cut-off criteria. This is extremely valuable information, because knowing the lower bound for plausible values of fit indices just due to random variation allows the creation of guidelines for use of fit statistics. This is essential because many fit indices do not have an easily derivable analytic distribution.

The use of acceptance rates alone does not tell the whole picture of the ability of fit indices to identify the correctly specified model. Furthermore, misspecified models may also have values common fit statistics that satisfy the cut-off criteria (e.g., the CFI $>$ .95). Thus, I conducted receiver-operating-characteristic (ROC) analysis to study the relationship between the sensitivity and specificity of cut-off criteria of each fit statistic. Sensitivity refers to how often the cut-off criteria correctly identifies the correct model, which can be optimized by making the criteria liberal and easy to meet. Specificity refers to how often only the correct model is identified, which can be optimized by making the criteria as conservative as possible and difficult to meet. A trade-off between these two criteria is therefore necessary as both are aimed at being optimized. The trade-off can be quantified by calculating the area under the curve (AUC) across a range of cut-off criteria.

In addition to the calculation of AUC, the partial-AUC (pAUC) was calculated. The pAUC represented the area under the curve for a range of the curve. That is, the classification quality is not only interest across all possible combinations of specificity and sensitivity, but on how sensitive an identifier is that is highly specific. This means that we restrict the range of our ROC curve to only the most specific (i.e., specificity of .9-1 or some other range of interest) to see how much area under this region is capture by our classifier. In this paper, I only reported standardized pAUCs so that the scale would be the same as tradition AUCs where 1 is perfect discrimination and .5 is no discrimination.

*Simulation Summary*

For each condition, 500 replications were used leading to 36,000 datasets being generated. Additionally, for each dataset simulated, four model specifications by three estimators were tested. Therefore, 12 models were estimated for dataset with a total of 432,000 models estimated.

The specific outcomes for this study are

(1) the distribution of fit statistics (CFI, TLI, RMSEA, SRMRW, and SRMRB) across design factors,

(2) influence of manipulated design factors ($\omega^2$ and partial-$\omega^2$),

(3) performance of fit statistic cut-off criteria. Tested with ROC analysis and acceptance rates across conditions and estimators.

In addition to the above outcomes, convergence rates of fitted models across conditions were reported along with any other estimation issues.

# CHAPTER FOUR

## Simulation Results

### *Convergence and Admissibility*

The entirety of the results of this simulation study are available online (Padgett, 2019). For each of the 72 conditions, 500 replications per cell were estimated. However, in Monte Carlo simulation studies the rates of convergence and the proportion of admissible solutions must be checked prior to examination of results. Convergence is the number of times M*plus* returned a result whereas the admissibility is the number of properly estimated solutions (i.e., no negative variances or impossible values). In keeping with previous simulation studies (DiStefano & Morgan, 2014; Flora & Curran, 2004; Yang-Wallentin, Jöreskog, & Luo, 2010) these improper solutions were removed from downstream analyses because they do not provide useful information (Forero et al., 2009). The remaining models were used in all subsequent analyses for summarizing model fit statistics and for graphic displays.

All three estimators converged at high rates across conditions. The ULSMV estimator exhibited the lowest convergence rates, but these rates were still about 97% of all models, see Table 4.1 for more details. A more fine grained breakdown of convergence rates across all 72 conditions, models and estimators was provided in Appendix C Table C.1. The lowest rates of convergence (.842-.866) were found when the level-1 model was misspecified (M1), sample sizes were low ($N_1 = 5$, $N_2 = 30$), and the estimator was WLSMV. The ULSMV and WLSMV estimators had similar rates of convergence across most conditions, though. Lower rates of convergence also tended to occur when the latent variable ICC ($ICC_L$) was low. However, convergence was less of an issue than properly estimated solutions (i.e., admissibility).

Table 4.1

*Proportion of converged and admissible solutions*

| | Convergence | | | Admissibility | | |
|---|---|---|---|---|---|---|
| Model | MLR | ULSMV | WLSMV | MLR | ULSMV | WLSMV |
| C | 1.000 | 0.999 | 1.000 | 0.834 | 0.765 | 0.722 |
| M1 | 0.982 | 0.974 | 0.965 | 0.731 | 0.610 | 0.524 |
| M2 | 1.000 | 0.988 | 1.000 | 0.832 | 0.772 | 0.733 |
| M12 | 1.000 | 0.985 | 0.992 | 0.830 | 0.641 | 0.617 |

*Note.* Each cell in the Convergence side of the table is out of 36,000 possible fitted models (500 rep. $\times$ 72 conditions). Each cell size in the Admissibility side of the table is conditional on the convergence rate for the corresponding cell. For example, for correctly specified models (C) under MLR estimator had a total of 30,024 ($32,000 \times 1 \times 0.834$) models with usable information. MLR = maximum likelihood with robust standard errors; ULSMV = unweighted least squares mean and variance adjusted; WLSMV = weighted least squares mean and variance adjusted; C = correctly specified models; M1 = model misspecified at level-1; M2 = model misspecified at level-2; and M12 = model misspecifed at both levels.

Across the model specifications and estimators, the highest rates of admissibility solutions were below 85%. The WLSMV estimator had the lowest rates of admissible solutions across model specifications (see Table 4.1). This is in line with previous simulations studies of WLSMV (DiStefano & Morgan, 2014; Forero et al., 2009). Models that were only misspecified at level-1 had the lowest rates of proper solutions. The rates of proper solutions for each cell of this design is broken down in Table C.2. When the level-1 model was misspecified numerous conditions resulted in zero usable cases. For example, using ULSMV or WLSMV resulted in zero usable cases with then $\text{ICC}_L = .5$ and $\text{ICC}_O = .1$ for most sample sizes. Interestingly, the rates of admissible solutions increased for correctly specified models as sample size ($\text{N}_1$ and $\text{N}_2$) increased, but decreased as sample size increased for incorrectly specified models at level-1. This trend was not observed for incorrectly specified models at level-2 or misspecified models at both levels.

The high rates of inadmissible solutions across some study conditions is concerning, but these rate are not unusual in Monte Carlo simulation studies with categorical CFA (DiStefano & Morgan, 2014; Forero et al., 2009; Yang-Wallentin et al., 2010). For example, DiStefano and Morgan (2014) reported zero usable replications under one condition of low sample size with skewed indicator distribution. In this study, the number of usable replications varied across conditions and model specification, and the remaining analyses may be limited by this reduction in sample size. For example, ROC analyses may not have sufficient number of cases in the misspecified models to be able to adequately differentiate between correct and incorrect specification.

### *Effect of Study Design Factors*

The effects of the study design on the distribution of fit statistics was investigated using an experimental design framework (i.e., ANOVA). Analyzing Monte Carlo simulation experiments using an ANOVA approached is recommended for distilling the important contributions of each chosen design factor on the outcomes of interest. In this study, six design factors were of interest; namely, the number of level-1 sampling units ($\mathrm{N}_1$), the number of groups sampled ($\mathrm{N}_2$), the ICC for observed variables ($\mathrm{ICC}_O$), the ICC for latent variables ($\mathrm{ICC}_L$), the model specification, and the model estimator. A full factorial ANOVA was not investigated for each fit statistic for two major reasons. First, the cells of the full factorial would be sparse under many conditions. The number of usable cases in multiple cells of the design were zero or less than 5% of the intended sample size of 500 per cell (see Table C.2). Therefore, the full factorial would not be able to properly estimate the effect of some higher order interactions. Secondly, higher order interactions do not provide a practically useful interpretation that would help guide researchers on the use of the chosen fit statistics. The bivariate interactions of all design factors provides a straightforward investigation of the impact of these factors on the distribution of each fit statistic.

The results of the ANOVA for each fit statistic are reported in Table 4.2. The results are reported as the partial-$\omega^2$ effect size estimates. The meaning of each value can be interpreted as follows for the $N_1$ (number of level-1 units) effect. For CFI, 8.4% of the variability in observed scores can be attributed to the number of level-1 units ($N_1$) that were sampled per group after controlling for all other design factors (i.e., level-2 sample size, observed and latent ICCs, model specification, and estimator).

Table 4.2

*Summary of ANOVA by effect size estimates with partial-$\omega^2$*

| Effect | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|
| $N_1$ | 0.084 | 0.084 | 0.163 | 0.658 | 0.179 |
| $N_2$ | 0.028 | 0.029 | 0.023 | 0.666 | 0.581 |
| $ICC_O$ | 0.011 | 0.011 | 0.140 | 0.159 | 0.364 |
| $ICC_L$ | 0.010 | 0.010 | 0.008 | 0.042 | 0.138 |
| Model | 0.418 | 0.417 | 0.600 | 0.770 | 0.078 |
| Estimator | 0.051 | 0.051 | 0.261 | 0.449 | 0.191 |
| $N_1$:$N_2$ | 0.076 | 0.076 | 0.060 | 0.281 | 0.023 |
| $N_1$:$ICC_O$ | 0.010 | 0.011 | 0.005 | 0.000 | 0.140 |
| $N_1$:$ICC_L$ | 0.001 | 0.001 | 0.009 | 0.005 | 0.043 |
| $N_1$:Model | 0.002 | 0.002 | 0.010 | 0.113 | 0.002 |
| $N_1$:Estimator | 0.004 | 0.004 | 0.004 | 0.015 | 0.000 |
| $N_2$:$ICC_O$ | 0.003 | 0.003 | 0.020 | 0.010 | 0.018 |
| $N_2$:$ICC_L$ | 0.012 | 0.012 | 0.005 | 0.007 | 0.058 |
| $N_2$:Model | 0.007 | 0.007 | 0.043 | 0.097 | 0.005 |
| $N_2$:Estimator | 0.072 | 0.072 | 0.149 | 0.074 | 0.010 |
| $ICC_O$:$ICC_L$ | 0.007 | 0.007 | 0.008 | 0.001 | 0.117 |
| $ICC_O$:Model | 0.016 | 0.016 | 0.062 | 0.006 | 0.023 |
| $ICC_O$:Estimator | 0.014 | 0.014 | 0.042 | 0.085 | 0.001 |
| $ICC_L$:Model | 0.040 | 0.040 | 0.068 | 0.017 | 0.050 |
| $ICC_L$:Estimator | 0.029 | 0.029 | 0.002 | 0.108 | 0.008 |
| Model:Estimator | 0.043 | 0.043 | 0.065 | 0.026 | 0.004 |

*Note.* The meaning of each value can be interpreted as follows. For example, for the effect of level-1 sample size ($N_1$), 8.4% of the variability in observed CFI scores can be attributed to the number of level-1 units were sampled per group after controlling for all other design factors (i.e., $N_2$, $ICC_O$, $ICC_L$, and bivariate interactions).

The results for CFI and TLI were nearly identical. The most influential factors on observed distribution of CFI and TLI is the model specification (41.7 and 41.8% for CFI and TLI, respectively), and the next highest estimated effects were level-1 sample size ($N_1$) that came out to be approximately equal at 8.4%.

In this simulation experiment, five of the design factors accounted for at least 10% of the variability in RMSEA scores after controlling for all other effects. The five factors are level-1 sample size ($N_1$), observed variable ICC ($ICC_O$), latent variable ICC ($ICC_L$), model specification, model estimator, and the interaction between number of groups ($N_2$) and estimator. These effects accounted for between 14% to 60% of variability in RMSEA after controlling for all other effects. The interaction effect of $N_2$ and estimator is especially interesting given that $N_2$ itself did not account for much variation in RMSEA scores (i.e., approximately 2.3%). This interaction effect would mean that the influence of the number of groups on the distribution of RMSEA depends on which estimator used. The dependent effect of number of groups is that for MLR, as the number of groups increases the distribution of RMSEA becomes less variable with lower values on average. For ULSMV and WLSMV, as the number of groups increases the distribution also becomes less variable. However, for these least-squares estimators, the values of RMSEA either do not change on average or slightly increase on average as level-2 sample size increases.

Five other interaction effects accounting for more than 10% of variability were also found across the SRMRW and SRMRB statistics. The model specification accounted for a large portion of variability (77%) for SRMRW but only a small portion (7.8%) for SRMRB. This differential effect is consistent with Figure 4.4 and Figure 4.5. Because SRMRB was not heavily influenced by model specification, which was the point of investigating SRMRB for model diagnostic utility, I focused the remainder this discussion on SRMRW. Within this study, three interaction effects account for greater than 10% of variability in SRMRW values after controlling for all other factors. The interactions were 1) effect of number of level-2 units is depends on the number of level-1 units sampled (partial-$\omega^2$ = .281), 2) the effect of number of level-1 unit depends on model specification (partial-$\omega^2$ = .113), and 3) the effect of $\text{ICC}_L$ depends on the model estimator (partial-$\omega^2$ = .108). Additional information on the effects of these design factor are given in Table B.1, similar trends are present. But this table reports $\omega^2$, which does not control for "off" effects.

*Distribution of Fit Statistics*

The distributions of fit statistics across model specification and estimators are shown in Figures 4.1, 4.2, 4.3, 4.4, and 4.5. These figures demonstrate the variability in the distribution of each commonly used fit statistic varies across model specification and estimator. The marginal distributions of all fit statistics across all 72 conditions are reported in Appendix E. In the results that follow, I mainly focused on the how the distribution of fit statistics varied (or lack of variation) across estimators and model specification. The estimator and model effect were the focus because these two effects accounted for the most variability in all fit statistics.

The distribution of CFI seems to have two major features that vary across model specifications and model estimators (see Figure 4.1). The means of CFI values differed across model specification. The most obvious differences in means of CFI scores occurs when the level-1 model is misspecified. This trend is seen across all estimators. Secondly, the distribution of CFI changes in variability mostly due to which estimator was used. For example, the range of all CFI scores under WLSMV is only about .6 to 1 while for MLR the range spans from 0 to 1 (i.e., the full range of potential CFI scores). These results were also found only using one type of misspecification (i.e., the omission of one cross loading). MLR also yielded the lowest value for a correctly specified model (CFI=.214) and this occured in condition 1 where sample sizes were

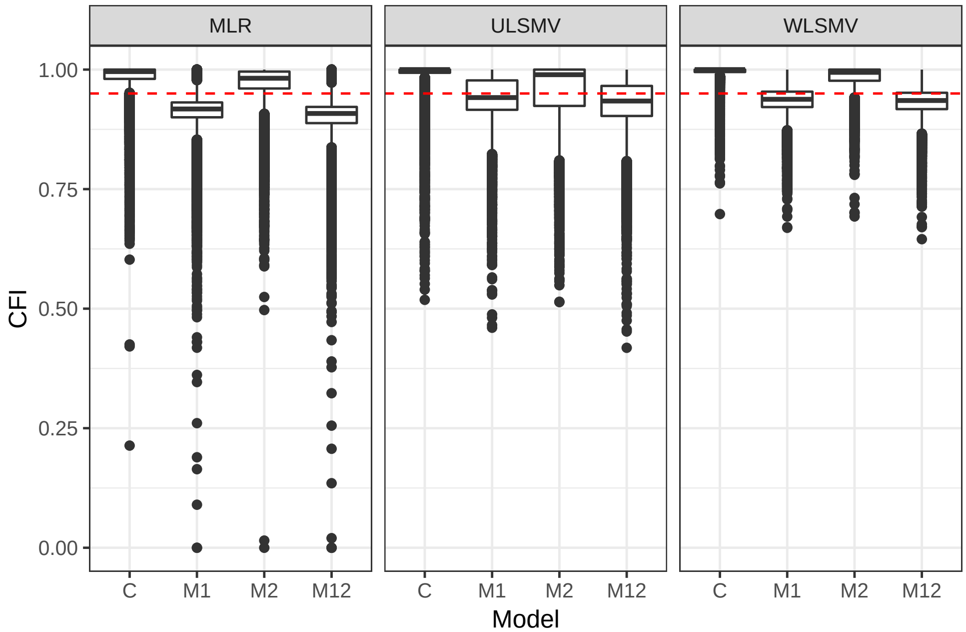



*Figure 4.1.* Distribution of CFI across estimated models and estimators.
*Note.* Dashed (red) line represents the Hu & Benter (1999) commonly reported cut-off for CFI at .95.

smallest ($N_1 = 5$, $N_2 = 30$) and ICCs were low ($ICC_L = .1, ICC_O = .1$). This trend of unequal variability is found primarily between estimators and not models.

For the TLI statistic, a some similar trends compared to CFI can be seem when looking at the distribution across model specification and estimators (see Figure 4.2). The trend similar between CFI and TLI is especially similar in terms of how variability in TLI values changes across estimators. The WLSMV had the least variability in TLI estimates. Another similarity between CFI and TLI is the potential for extremely low estimates of TLI even for correctly specified models (e.g., see Figure 4.2 where many values resulted in values less than .75 across estimators). In contrast to CFI, the TLI scores under MLR estimation yielded greater separation in means between model specification. The observed difference between the Q3 for model C and Q1 for model M1 are greater for TLI and CFI. Lastly for TLI, the IQR for models M1, M2, and

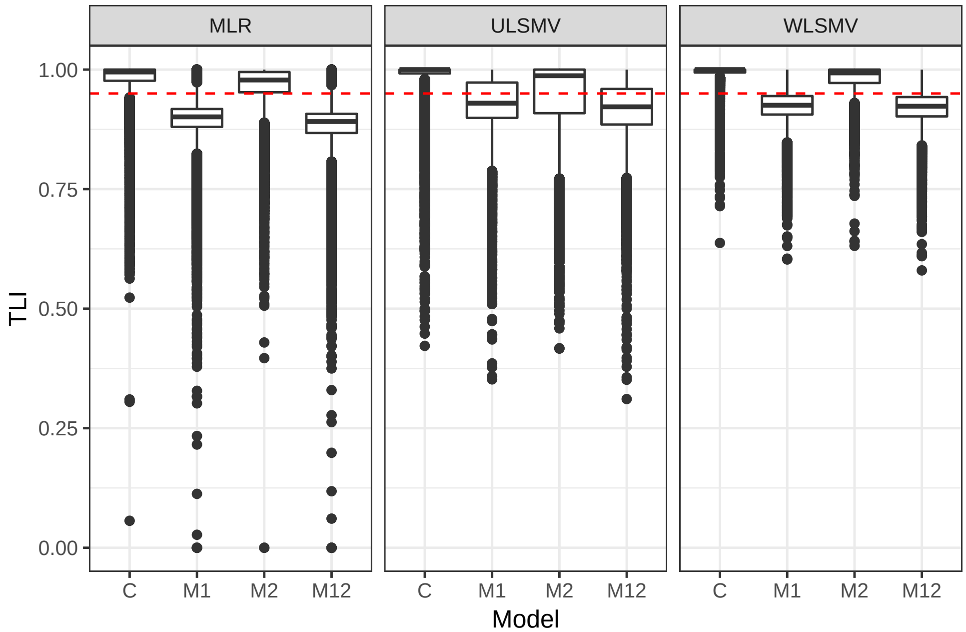



*Figure 4.2.* Distribution of TLI across estimated models and estimators.
*Note.* Dashed (red) line represents the Hu & Benter (1999) commonly reported cut-off for TLI at .95.

M12 is greatest under the ULSMV estimator, meaning that the potential range for likely values is greatest with the ULSMV despite MLR returning more extreme cases.

The reader is cautioned on the interpretation of the TLI scores because TLI is non-normed, meaning the potential range of scores is between positive and negative infinity. TLI doesn't have the range specification that CFI does. In this project, all TLI scores greater than one were rescaled to 1, as is commonly done in practice. Also, all values less than zero were scaled to zero.

The distribution for RMSEA is shown across models and estimators in Figure 4.3. One aspect of the distribution of RMSEA that immediately stands out is that for ULSMV and WLSMV nearly all estimates fall below the commonly used cut-off of .06. This trend is even more apparent when one looks at the distribution of RMSEA across all conditions irrespective model and estimator (see Figure E.3 in the Appendices). Additionally, across all models and estimators, the IQRs for the

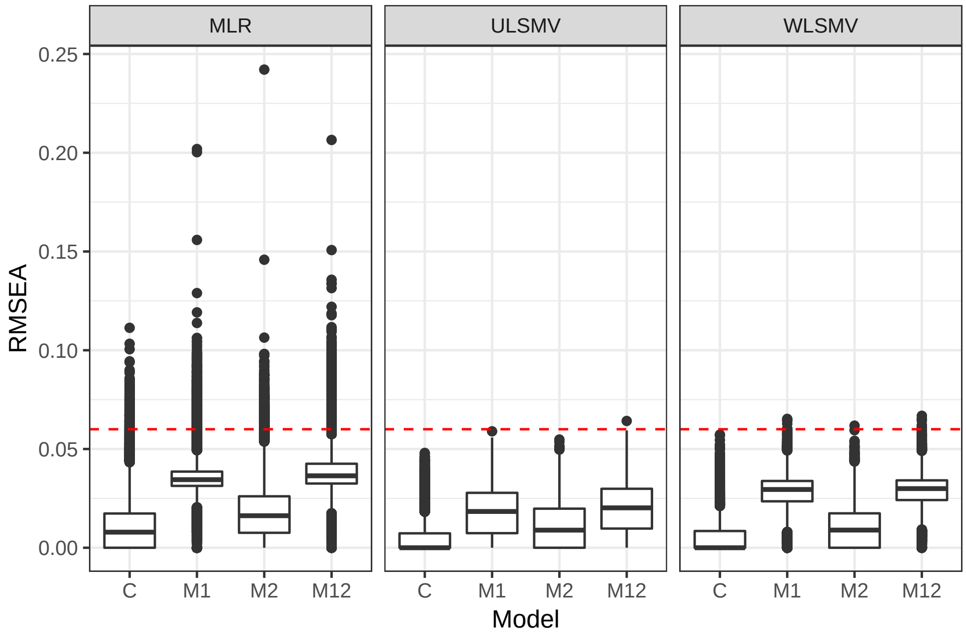



*Figure 4.3.* Distribution of RMSEA across estimated models and estimators.
*Note.* Dashed (red) line represents the Hu & Benter (1999) commonly reported cut-off for RMSEA at .06.

distribution of RMSEA were below the commonly used cut-off. However, despite the astoundingly low estimates across all conditions and estimators, the estimates for model C were lower on average. The largest changes in RMSEA appear to due to misspecification of the level-1 factor structure.

The distribution for SRMRW is shown across models and estimators in Figure 4.4. SRMRW is designed to help detect when the level-1 (Within) covariance structure is misspecified (Hsu et al., 2015). Two major features stand out for SRMRW. First, there appears to be a clear discrimination between models that have a correctly specified level-1 model and those models that do not (similar to CFI, TLI, and RMSEA). This can be seem by apparent differences between the IQRs across models, and this pattern holds across estimators. Despite similar patterns across estimators the ULSMV yielded the most variable estimates on average across model specification, but MLR resulted in the highest values for incorrect model specification. Secondly, no

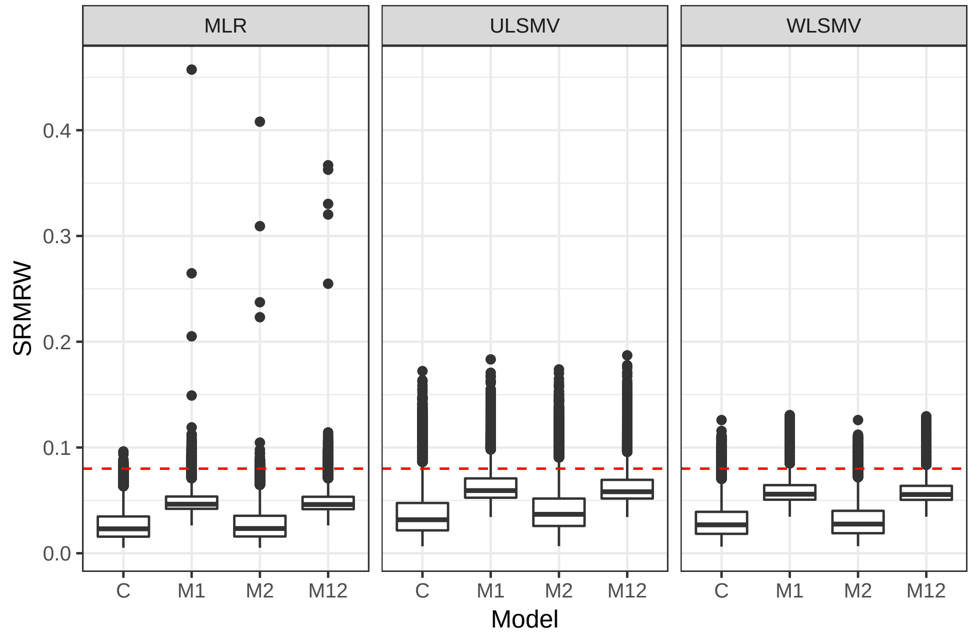



*Figure 4.4.* Distribution of SRMRW across estimated models and estimators.
*Note.* Dashed (red) line represents the Hu & Benter (1999) commonly reported cut-off for SRMR at .08.

obvious difference between model C and model M2 can be discerned at first glance. This lack if discrimination between model C and moel M2 is expected given that SRMRW approximately estimates the standardized average distance between the observed within covariance matrix and the model implied covariance matrix (see the Methods section on SRMRW for more information on the computation).

The distribution for SRMRB is shown across models and estimators in Figure 4.5. SRMRB is designed to help detect when the level-2 (Between) covariance structure is misspecified (Hsu et al., 2015). One major feature of the distributions of SRMRB stands out. That one feature is little variation we observed between models and estimators. Prior literature suggested that SRMRB would perform best at discriminating between a correct and incorrect level-2 covariance structure specification. However, based on Figure 4.5 we could not identify a clear pattern to SRMRB across

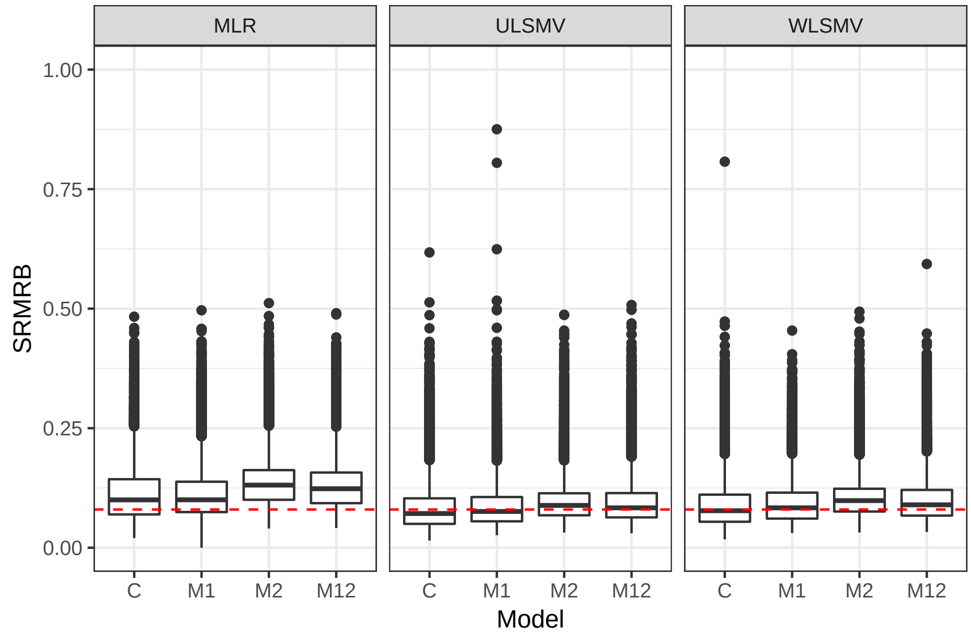



*Figure 4.5.* Distribution of SRMRB across estimated models and estimators.
*Note.* Dashed (red) line represents the Hu & Benter (1999) commonly reported cut-off for SRMR at .08. We fixed the range to be 0,1 for viewing purposes; however, the max value we observed was 8.49.

models and estimators. There is a slight visible increase in thevalues of SRMRB under estimation with MLR, but this increase is so small that there practically no difference.

Aside from the small amounts of variability observed among model specifications, we noticed one peculiar attribute of the distribution of SRMRB in our study. Some excessively large values were observed in Conditions 1 and 7. In condition 7, where the number of groups ($N_2 = 30$) with units per group ($N_1 = 10$) with low ICCs ($ICC_L = .1, ICC_O = .1$), the max value observed was 8.49 for an apparently admissible solution occurred under WLSMV, Model M1, and according to CFI, TLI, and RMSEA, this model fits perfectly, and the $\chi^2$ test of goodness of fit was not rejected. The highest value observed for Model C was 2.54, in the same condition 7. Both of these large estimates were based on the same data (Rep. 489) in condition 7, so it possible that these data simply did not provide enough variability in observed scores

Table 4.3

*Summary of fit statistics across model specification and estimators*

| Estimator | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|
| Model C | | | | | | |
| MLR | .820 | .981 (0.04) | .977 (0.05) | 0.012 (0.01) | 0.027 (0.02) | 0.113 (0.06) |
| ULSMV | .977 | .989 (0.03) | .987 (0.04) | 0.004 (0.01) | 0.037 (0.02) | 0.082 (0.05) |
| WLSMV | .975 | .994 (0.02) | .993 (0.02) | 0.005 (0.01) | 0.031 (0.02) | 0.088 (0.05) |
| Model M1 | | | | | | |
| MLR | .060 | .910 (0.05) | .892 (0.05) | 0.036 (0.01) | 0.049 (0.01) | 0.112 (0.05) |
| ULSMV | .478 | .940 (0.05) | .928 (0.06) | 0.018 (0.01) | 0.064 (0.02) | 0.087 (0.05) |
| WLSMV | .227 | .937 (0.03) | .924 (0.04) | 0.028 (0.01) | 0.059 (0.01) | 0.094 (0.08) |
| Model M2 | | | | | | |
| MLR | .523 | .970 (0.04) | .963 (0.05) | 0.018 (0.02) | 0.027 (0.02) | 0.137 (0.05) |
| ULSMV | .716 | .954 (0.07) | .945 (0.08) | 0.011 (0.01) | 0.041 (0.02) | 0.097 (0.04) |
| WLSMV | .751 | .985 (0.02) | .982 (0.03) | 0.010 (0.01) | 0.032 (0.02) | 0.104 (0.04) |
| Model M12 | | | | | | |
| MLR | .044 | .899 (0.05) | .880 (0.05) | 0.039 (0.01) | 0.049 (0.01) | 0.131 (0.05) |
| ULSMV | .424 | .929 (0.06) | .916 (0.07) | 0.020 (0.01) | 0.063 (0.02) | 0.095 (0.05) |
| WLSMV | .211 | .933 (0.03) | .921 (0.04) | 0.028 (0.01) | 0.059 (0.01) | 0.099 (0.04) |

*Note*. The reported values under $\chi^2$ are the proportion of times the $\chi^2$ test of goodness of fit reported a p-value greater than $\alpha = .05$. For remaining fit statistics (CFI, TLI, etc.) the reported values are the average value for that fit statistic where values in parentheses are the standard deviations.

for all the variances to be precisely estimated even though estimation terminated normally.

The empirical distribution of each fit statistics were shown in Figures 4.1-4.5. Next, the summary statistics of each fit index across model specification and estimators are reported in Table 4.3. Additional tables for the summary of the fit statistics across all conditions, models, and estimators are reported in Appendix D.

The summary of the $\chi^2$ test of goodness of fit is included in Table 4.3. For correctly specified models, the reported value should be approximately .95. The value should be .95 because this is the proportion of times the $\chi^2$ test resulted in a non-significant p-value, meaning that $p > \alpha = .05$. For all incorrectly specified models, this value should be close to .05, the Type I error rate. Under correctly specified models estimated by ULSMV and WLSMV, the $\chi^2$ test resulted in a higher proportion of non-significant hypothesis tests. However, the $\chi^2$ test also did return significant results when there should have been (i.e., models M1, M2, and M12). The $\chi^2$ test performed better when estimated under MLR at least in terms of error rates for incorrectly specified models.

*ROC Analyses Results*

One of the defining features of using ROC analysis is the creation of ROC curves. A ROC curve helps create a visual representation for the quality of classification based on a systematically varying cutoff criteria. In this case, the cutoff criteria is the value of the fit statistic that differentiate between a correctly specified model and an incorrectly specified model. In MCFA models, there are three ways a model can be misspecified (see Methods Summary of Models Fitted). Three ROC analyses were initially performed for detecting 1) any misspecification (Model C versus Model M1-M2-M12); 2) mispecification at level-1 (Model C-M2 versus Model M1-M12); and 3) mispecification at level-2 (Model C-M1 versus Model M2-M12).

The results for ROC analyses across conditions irrespective of estimator are shown in Figure 4.6 and Table 4.4. The curves in Figure 4.6 help describe how strongly each fit index discriminates between the correctly specified model and the misspecified models. Irrespective of conditions and estimators, CFI, TLI, and RMSEA appear to perform approximately equally. The first (left) panel of Figure 4.6 shows how the predictive value for detecting the correct model out three specifications varies across fit indices. SRMRW and SRMRB have a poor discriminate curve. The discrimination of SRMRW and SRMRB can also be seen as inadequate or at least under performing compared to CFI, TLI, and RMSEA based on the AUC value for each index reported

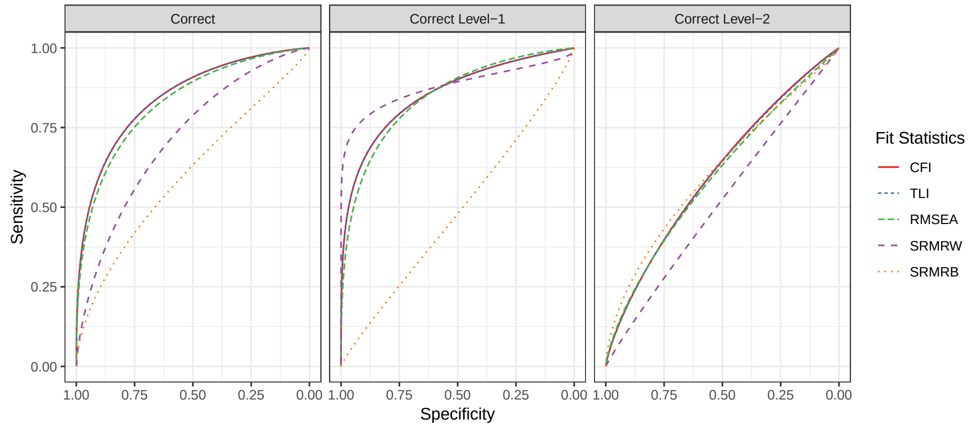



*Figure 4.6.* ROC analysis curves for identifying correctly specified models.
*Note.* The Correct panel reflects the ROC analyses for identifying the data generating model (i.e., the completely correct specification) over the three types of misspecification. The Correct Level-1 ROC analyses reflect the comparison between models with a correctly specified level-1 model versus models that have any misspefication of the level-2 model (i.e., Models C and M2 vs. Models M1 and M12). The last comparison of Correct Level-2 was among models with a correctly specified level-2 model versus models with any level-1 misspecification (i.e., Models C and M1 versus Models M2 and M12).

in Table 4.4. SRMRB has a poor AUC of .598, though the performance of SRMRW appears only slightly behind that of CFI, TLI, and RMSEA at.742.

SRMRW had a lower AUC than CFI, TLI, and RMSEA for detecting any model misspecification detection. However, SRMRW provided the sharpest transition for detecting a level-1 model misspecification (see middle panel of Figure 4.6). For this middle panel, SRMRB has essentially zero predictive power of level-1 misspecification, because the curve is a a straight line along the diagonal. In the right most panel, SRMRB performs as well as CFI, TLI, and RMSEA at predicting level-2 misspecification. Although this predictive performance is barely greater than chance for all indices.

Despite the apparent poor performance of these fit statistics at detecting level-2 misspecification, there is at least potential for identifying whether there is any misspecification in a model. The results presented in Table 4.4 also provide the summary

Table 4.4

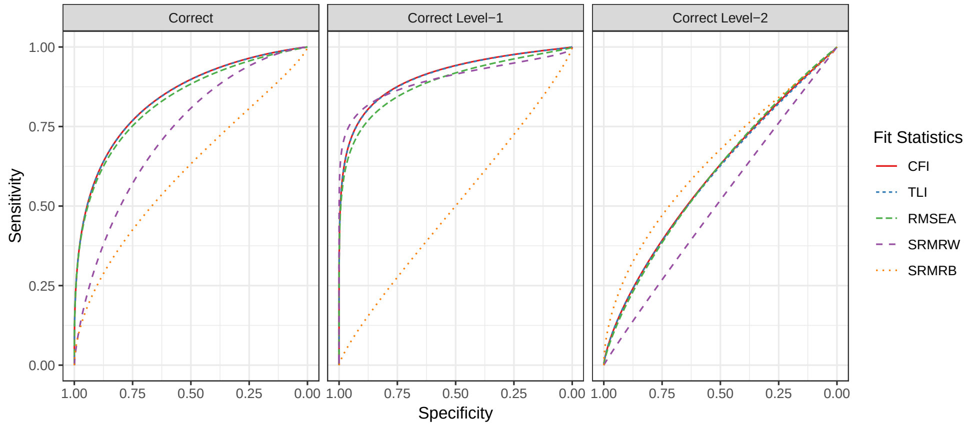
*Overall ROC analysis AUCs and optimized threshold*

| Index | AUC[1] | pAUC[2] | Threshold[3] | Specificity | Sensitivity |
|---|---|---|---|---|---|
| CFI | 0.816 | 0.565 | 0.977 | 0.702 | 0.855 |
| TLI | 0.815 | 0.565 | 0.972 | 0.702 | 0.855 |
| RMSEA | 0.803 | 0.565 | 0.015 | 0.685 | 0.829 |
| SRMRW | 0.742 | 0.551 | 0.038 | 0.728 | 0.723 |
| SRMRB | 0.598 | 0.557 | 0.067 | 0.804 | 0.352 |

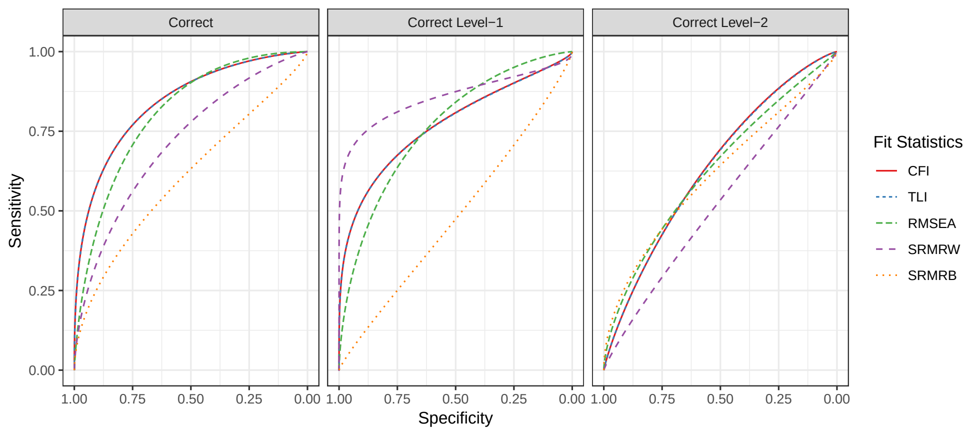
*Note.* [1] AUC = Area Under the Curve;
[2] pAUC = partial-AUC (specificity of .9-1); and
[3] Optimal-Threshold determined by best Specificity and Sensitivity.

information of the ROC curves given the left panel of Figure 4.6. The Overall optimal threshold for CFI irrespective of estimator is .977, with an AUC of .816 and pAUC of .565. These results break down to mean that, first, a cutoff value of .977 optimizes the specificity and sensitivity of classification a model as correctly specified over any type of misspecification. Secondly, an AUC of .816 means that there is approximately a .816 probability that a model with a CFI value above .977 is correctly specified. And lastly, that a pAUC of .565 represents the average sensitivity of correctly identifying the estimating model conditional on being highly specific (i.e., specificity greater than .9). That is, for CFI cut-offs that are highly discriminatory between correct and incorrect model, these cut-off criteria are not sensitive enough to identify the correct models.

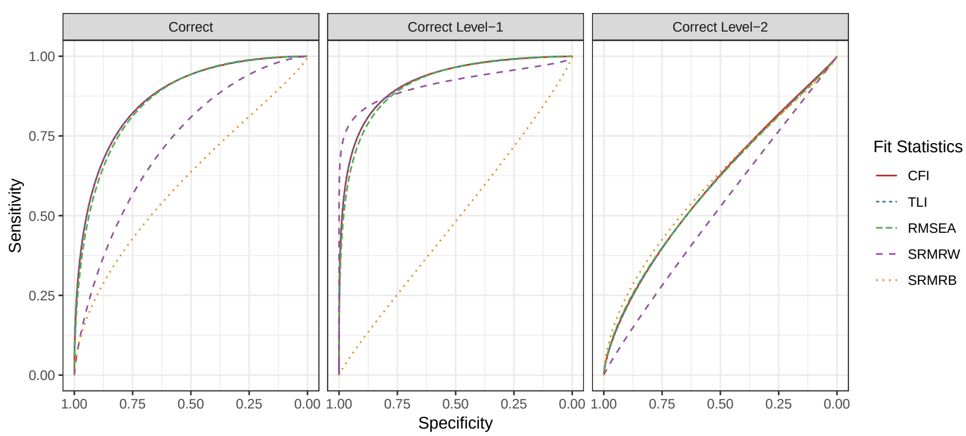
The results for ROC analyses across conditions for each estimator are shown in Figure 4.7. The overall trend across estimators was that a similar shape occurred between the three types of ROC analyses performed (Correct,Correct Level-1 vs. Correct Level-2), these three plots have a similar trend across estimators with only minor differences. That is, the estimator effect is difficult to detect from a coarse view of the ROC analyses. However, in Tables 4.5-4.7 subtle differences can be parsed a out across estimators. One of these estimators differences was found in the identification




*(a)* MLR Estimator




*(b)* ULSMV Estimator




*(c)* WLSMV Estimator

*Figure 4.7.* ROC analysis curves across estimators

of correctly specified models (left most panels of Figure 4.7). The WLSMV estimator yielded the sharpest curve for CFI, TLI, and RMSEA compared to MLR and ULSMV (Figures 4.7a-4.7b). Under model estimation with WLSMV, these statistics appear to performed best at identifying the completely correct model specification compared to any form of misspecification. For more information see Table 4.5, where the AUCs under MLR and WLSMV are almost identical.[1] Another subtle difference among these three estimators is the shape of the ROC curves under correctly specified level-1 models. There is evidence of some differences among the signal to noise ratios being examined, although the exact differences are difficult to interpret for Figure 4.7.

To parse out these subtle differences in estimator effects on model selection are the results reported in Tables 4.5-4.7. Because we used smoothing for plotting the ROC curves, the AUCs for WLSMV were slightly inflated.

The inflated AUCs of the smoothed ROC curves resulted in the apparent increase in classification quality of WLSMV found in Figure 4.7c. However, the AUCs for MLR and WLSMV were similar enough to potentially make the differences negligible. Note, we did not conduct any formal statistical tests to compare these AUCs, though these comparisons are possible for future investigations.

Similar to the overall ROC analyses, the SRMRB performed poorly at identifying the correctly specified models across estimators (AUCs = .595-.601, see Table 4.5). However, interestingly the resulting "best" threshold for SRMRB under estimation with MLR was 0.08, which is the Hu and Bentler (1999) commonly reported cutoff criteria. For SRMRB (and all other fit statistics), the pAUCs (i.e., AUC conditional on a specificity range of specificity .9-1) across estimators ranged from .541-.627. This poor range of pAUCs gives evidence that these fit statistics are inadequately sensi-

[1] The differences be figures and reported AUCs are due to plotting with the smoothed curves instead of the empirical distribution. This was accomplished with the "smooth()" function in pROC with default settings in order to make graphing the final figures easier. The "smoothed" AUCs are reported in the online supplement material with the code.

Table 4.5

*Completely correct specification ROC analysis AUCs and optimized threshold by estimator*

| Index | Estimator | AUC[1] | pAUC[2] | Threshold[3] | Specificity | Sensitivity |
|---|---|---|---|---|---|---|
| CFI | MLR | 0.841 | 0.627 | 0.967 | 0.746 | 0.834 |
| | ULSMV | 0.780 | 0.541 | 0.973 | 0.628 | 0.888 |
| | WLSMV | 0.836 | 0.571 | 0.982 | 0.682 | 0.896 |
| TLI | MLR | 0.840 | 0.627 | 0.961 | 0.750 | 0.830 |
| | ULSMV | 0.780 | 0.541 | 0.968 | 0.627 | 0.888 |
| | WLSMV | 0.836 | 0.571 | 0.979 | 0.682 | 0.896 |
| RMSEA | MLR | 0.830 | 0.625 | 0.024 | 0.735 | 0.843 |
| | ULSMV | 0.770 | 0.541 | 0.012 | 0.596 | 0.855 |
| | WLSMV | 0.832 | 0.571 | 0.014 | 0.689 | 0.872 |
| SRMRW | MLR | 0.754 | 0.551 | 0.036 | 0.720 | 0.773 |
| | ULSMV | 0.740 | 0.569 | 0.045 | 0.741 | 0.715 |
| | WLSMV | 0.750 | 0.547 | 0.044 | 0.671 | 0.813 |
| SRMRB | MLR | 0.600 | 0.563 | 0.083 | 0.800 | 0.366 |
| | ULSMV | 0.595 | 0.568 | 0.056 | 0.838 | 0.324 |
| | WLSMV | 0.601 | 0.563 | 0.062 | 0.822 | 0.345 |

*Note.* [1] AUC = Area Under the Curve; [2] pAUC = partial-AUC (specificity of .9-1); and [3] Optimal-Threshold determined by best Specificity and Sensitivity.

tive to differences in model specification when being highly discriminant in model selection is warranted.

In general, the discrimination of these common statistics appears low, but we still need to make a decision as to the adequacy of model fit. To help with this judgement, many researchers use Hu and Bentler (1999) cut-off criteria for CFI, TLI, RMSEA, and SRMRW. However, these values were not recovered in our scenario based on these ROC analyses for completely correct model selection. For CFI, the lowest cutoff across the three estimators was .967 (so essentially .97) where WLSMV returned the highest optimal threshold of .98. For TLI, a similar trend was found as for CFI, except that optimal thresholds were .961-.979. RMSEA thresholds were found to be much lower, which was likely due to sample size required for these data, especially given that we already found RMSEA to be sensitive to level-1 sample size (see Table 4.2). SRMRW was yielded lower overall classification quality (AUC .74-75). The thresholds were also lower, but the sensitivity at this threshold was also a little lower than for CFI, TLI, and RMSEA. The SRMRB performed poorly enough to where any threshold is essentially useless in this case.

Beyond complete model selection with these fit indices, there is potential that these statistics can help identify where part of a model is correctly specified. The ROC analyses for classifying models as having a correct level-1 model specification are reported in Table 4.6.

Table 4.6

*Correct level-1 specification ROC analysis AUCs and optimized threshold by estimator*

| Index | Estimator | AUC[1] | pAUC[2] | Threshold[3] | Specificity | Sensitivity |
|---|---|---|---|---|---|---|
| CFI | MLR | 0.925 | 0.888 | **0.950** | 0.944 | 0.854 |
| | ULSMV | 0.737 | 0.543 | 0.974 | 0.760 | 0.735 |
| | WLSMV | 0.924 | 0.756 | 0.968 | 0.875 | 0.886 |
| TLI | MLR | 0.924 | 0.887 | 0.940 | 0.944 | 0.854 |
| | ULSMV | 0.736 | 0.543 | 0.968 | 0.759 | 0.736 |
| | WLSMV | 0.923 | 0.756 | 0.962 | 0.874 | 0.886 |
| RMSEA | MLR | 0.899 | 0.868 | 0.027 | 0.939 | 0.817 |
| | ULSMV | 0.745 | 0.543 | 0.014 | 0.670 | 0.748 |
| | WLSMV | 0.916 | 0.731 | 0.020 | 0.840 | 0.880 |
| SRMRW | MLR | 0.884 | 0.885 | 0.037 | 0.971 | 0.776 |
| | ULSMV | 0.838 | 0.844 | 0.046 | 0.971 | 0.695 |
| | WLSMV | 0.902 | 0.905 | 0.044 | 0.978 | 0.809 |
| SRMRB | MLR | 0.513 | 0.504 | 0.134 | 0.651 | 0.388 |
| | ULSMV | 0.494 | NA[4] | 0.081 | 0.516 | 0.497 |
| | WLSMV | 0.500 | NA[4] | 0.084 | 0.473 | 0.550 |

*Note.* The CFI optimal threshold is bolded because the value is the exact same as the Hu & Bentler (1999) recommended criteria. [1] AUC = Area Under the Curve;
[2] pAUC = partial-AUC (specificity of .9-1);
[3] Optimal-Threshold determined by best Specificity and Sensitivity; and
[4] pAUC not computed due to error in the estimation.

The AUCs tended to be higher than from the previous ROC analyses across all fit statistics except SRMRB. The pAUCs were also much better across fit statistics.

This lends evidence that these fit statistics are not only sensitive across a wide range of specificities (AUCs), but also provide discrimination and identification of correctly specified level-1 models. However, when models were estimated with ULSMV, this trend was only observed for SRMRW. The index with the highest AUC (.925) was CFI, particularly when models were estimated with MLR. Interestingly, the optimal threshold found for CFI under MLR was the commonly used cutoff of .95 (see Table 4.6). Other commonly used thresholds were not found for any other statistic or estimator.

The ROC analyses for classifying models as having a correct level-2 model specification are reported in Table 4.7. The statistic that performed "best" was SRMR, but, the AUCs only range from .602-.640. These statistics do not appear to help discriminate among incorrect and correct level-2 models. This is contrary to some previous simulation studies that found SRMRB to be useful in identifying level-2 misfit (Hsu et al., 2015).

### *Acceptance Rates for Correctly Specified Models*

Acceptance rates are the proportion of times the correctly specified model has a value on the fit index at or above (or below depending on the statistic) a given value. A feature of an "optimal" threshold/cut-off value would be if the acceptance rate was at least 95% of all observed values for correctly specified models. For each of the five commonly used fit statistics investigated, the acceptance rate based on varying cut-off criteria are reported in Table 4.8. The acceptance rates are separated into two chunks, in the first lays CFI and TLI because they are scaled 0-1 with 1 being optimal. In the second lays RMSEA, SRMRW, and SRMRB because they are scales such that lower values are optimal. The reported hit rates in Table 4.8 are interpreted as the proportion of admissible solutions that have a fit value within the range of the cut-off criteria. For example, in the row for CFI, the value reported under the cutoff .91 is

Table 4.7

*Correct level-2 specification ROC analysis AUCs and optimized threshold by estimator*

| Index | Estimator | AUC[1] | pAUC[2] | Threshold[3] | Specificity | Sensitivity |
|---|---|---|---|---|---|---|
| CFI | MLR | 0.595 | 0.536 | 0.989 | 0.803 | 0.353 |
| | ULSMV | 0.630 | 0.518 | 0.971 | 0.574 | 0.624 |
| | WLSMV | 0.591 | 0.521 | 0.995 | 0.718 | 0.460 |
| TLI | MLR | 0.592 | 0.536 | 0.987 | 0.803 | 0.353 |
| | ULSMV | 0.629 | 0.518 | 0.966 | 0.573 | 0.624 |
| | WLSMV | 0.589 | 0.521 | 0.994 | 0.718 | 0.460 |
| RMSEA | MLR | 0.593 | 0.535 | 0.012 | 0.808 | 0.340 |
| | ULSMV | 0.615 | 0.518 | 0.007 | 0.673 | 0.515 |
| | WLSMV | 0.589 | 0.521 | 0.008 | 0.728 | 0.450 |
| SRMRW | MLR | 0.514 | 0.503 | 0.031 | 0.660 | 0.373 |
| | ULSMV | 0.525 | 0.511 | 0.032 | 0.786 | 0.283 |
| | WLSMV | 0.521 | 0.506 | 0.033 | 0.651 | 0.390 |
| SRMRB | MLR | 0.640 | 0.573 | 0.098 | 0.737 | 0.487 |
| | ULSMV | 0.603 | 0.571 | 0.058 | 0.845 | 0.329 |
| | WLSMV | 0.602 | 0.559 | 0.067 | 0.799 | 0.371 |

*Note.* [1] AUC = Area Under the Curve; [2] pAUC = partial-AUC (specificity of .9-1); and [3] Optimal-Threshold determined by best Specificity and Sensitivity.

interpreted as 97.1% of admissible solutions for correctly specified models haing CFI values at or above .91. The remaining values are interpreted analogously.

To obtain an acceptance rate of approx .95, the threshold would as low as .93 for CFI and .92 for TLI. For the second chunk of criteria, the cut-off would be .03 for RMSEA, .07 for SRMRW, and some value greater than .09 for SRMRB. The only statistic to have somewhat converging evidence between acceptance rates and the ROC analyses for an "optimal" cuf-off criteria is RMSEA. From the acceptance rates, an optimal threshold may be approximately .03 as the upper limit. Based on the ROC analysis reported in Table 4.4, the cut-off would be approx half at .015.

The global cut-offs discussed above do not account for the estimator effects that are present in these results. The acceptance rates by estimators are presented in Table 4.9. For CFI and TLI, a consistent trend emerged that for the same cut-off WLSMV yielded the highest acceptance rate, ULSMV had the middle value, and MLR had the lowest acceptance rate. This lends more evidence that using the same cut-off criteria across estimators may not be optimal given the differences in acceptance rates of each fit index across estimators. The changing acceptance rates would also lead to different rates of correctly identifying the correct model based on the same model conditional the estimator used.

Table 4.8

*Acceptance rates of fit statistics for select cut-off criteria*

| | Cut-off Value | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Index | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 |
| CFI | 0.971 | 0.965 | 0.957 | 0.947 | 0.933 | 0.914 | 0.886 | 0.838 |
| TLI | 0.961 | 0.954 | 0.945 | 0.933 | 0.918 | 0.896 | 0.865 | 0.814 |
| | Cut-off Value | | | | | | | |
| | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
| RMSEA | 0.895 | 0.958 | 0.981 | 0.991 | 0.997 | 0.999 | 1.000 | 1.000 |
| SRMRW | 0.306 | 0.577 | 0.747 | 0.851 | 0.913 | 0.952 | 0.975 | 0.987 |
| SRMRB | 0.001 | 0.017 | 0.069 | 0.165 | 0.278 | 0.383 | 0.477 | 0.557 |

*Note.* These data are subset to only the correctly specified model (Model C). Acceptance rates are the proportion of times the observed fit statistic fell within the range of the cutoff criteria. For CFI and TLI, the values represent the proportion of times the statistic was at least the cutoff value. For RMSEA, SRMRW, and SRMRB, the values represent the proportion of times the statistics was at most the cutoff value.

The differences that were observed between ULSMV and WLSMV for RMSEA were negligible. Both of the least-squares estimators however had higher acceptance rate than MLR across most cut-offs. The exception is the cut-offs of .08-.1, where the rates were nearly all a value of 1. The optimal cut-off for trying to get the acceptance rate approximately .95 would depend on the estimator used. For MLR, the cut-off value would be .04 while for ULSMV and WLSMV the cut-off value would be .02.

The cut-off for SRMRW based on the acceptance rate would also depend on the estimator used. For MLR, the cut-off was .06, which is higher than found by the ROC analysis (.036, see Table 4.5). Estimation with ULSMV would result in a cut-off of .09, which is higher than found by the ROC analysis (.045, see Table 4.5). Models estimated with WLSMV resulted in an optimal cut-off of .07, which is higher than found by the ROC analysis (.044, see Table 4.5). The cut-off is not consistent between methods for trying to triangulate an optimal threshold for SRMRW. For SRMRB, none of the cut-offs tested provided high acceptance rates across any estimators. So, a cut-off higher than .1 may potentially be useful for SRMRB.

Table 4.9

*Acceptance rates of fit statistics across estimators*

| | | Cutoff Value | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Index | Estimator | 0.9 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.97 | 0.98 | 0.99 |
| CFI | MLR | 0.958 | 0.949 | 0.939 | 0.926 | 0.910 | 0.889 | 0.861 | 0.821 | 0.755 | 0.631 |
| | ULSMV | 0.978 | 0.974 | 0.969 | 0.961 | 0.953 | 0.941 | 0.925 | 0.900 | 0.862 | 0.790 |
| | WLSMV | 0.996 | 0.994 | 0.991 | 0.989 | 0.984 | 0.976 | 0.964 | 0.946 | 0.909 | 0.831 |
| TLI | MLR | 0.942 | 0.932 | 0.920 | 0.906 | 0.889 | 0.867 | 0.835 | 0.791 | 0.723 | 0.600 |
| | ULSMV | 0.970 | 0.965 | 0.959 | 0.951 | 0.941 | 0.928 | 0.909 | 0.883 | 0.842 | 0.771 |
| | WLSMV | 0.992 | 0.990 | 0.987 | 0.983 | 0.976 | 0.966 | 0.953 | 0.930 | 0.890 | 0.808 |
| | | Cutoff Value | | | | | | | | | |
| | | 0.1 | 0.09 | 0.08 | 0.07 | 0.06 | 0.05 | 0.04 | 0.03 | 0.02 | 0.01 |
| RMSEA | MLR | 1.000 | 1.000 | 0.999 | 0.997 | 0.990 | 0.976 | 0.948 | 0.898 | 0.795 | 0.565 |
| | ULSMV | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 0.993 | 0.956 | 0.816 |
| | WLSMV | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 0.989 | 0.946 | 0.788 |
| SRMRW | MLR | 1.000 | 1.000 | 0.998 | 0.987 | 0.959 | 0.909 | 0.822 | 0.675 | 0.396 | 0.089 |
| | ULSMV | 0.982 | 0.968 | 0.943 | 0.907 | 0.853 | 0.775 | 0.653 | 0.460 | 0.211 | 0.017 |
| | WLSMV | 0.999 | 0.993 | 0.981 | 0.959 | 0.924 | 0.864 | 0.760 | 0.590 | 0.302 | 0.031 |
| SRMRB | MLR | 0.499 | 0.423 | 0.344 | 0.253 | 0.150 | 0.065 | 0.021 | 0.004 | 0.000 | 0.000 |
| | ULSMV | 0.729 | 0.659 | 0.578 | 0.483 | 0.377 | 0.251 | 0.123 | 0.035 | 0.003 | 0.000 |
| | WLSMV | 0.676 | 0.605 | 0.524 | 0.428 | 0.320 | 0.191 | 0.067 | 0.012 | 0.000 | 0.000 |

*Note.* These data are subset to only the correctly specified model (Model C). Acceptance rates are the proportion of times the observed fit statistic fell within the range of the cutoff criteria. For CFI and TLI, the values represent the proportion of times the statistic was at least the cutoff value. For RMSEA, SRMRW, and SRMRB, the values represent the proportion of times the statistics was at most the cutoff value.

# CHAPTER FIVE

## Discussion

Multilevel measurement models allow researchers to test hypotheses of complex phenomena using data from complex sampling methods. The complex organizational structure of educational and psychological data are particularly well suited for these types of multilevel measurement models; however, finding evidence for the validity of these measurement model specific is not well established. The specified measurement models, the use of common fit indices (e.g., CFI, TLI, RMSEA, SRMRW, SRMRB) should be interpreted with caution because the sensitivity of commonly used fit statistics to misspecification may depend on the estimator used. The influence of robust estimators on fit indices in complex settings is still considerably unknown (Asparouhov & Muthén, 2007; DiStefano & Morgan, 2014; Hsu, 2009; Navruz, 2016; Nestler, 2013). For example, DiStefano and Morgan (2014) found that the CFI and RMSEA were generally not sensitive to variation in number of response categories, nonnormality, and sample size, but slightly poorer estimates of fit were observed when number of categories was only two and the response distribution was skewed. In their study, the performance of CFI and RMSEA may be because models were specified correctly. The major difference in the current study is that the categorical CFA model is extended into the multilevel space.

In multilevel CFA with categorical data, CFI *and* TLI was generally only infuenced by misspecification of the level-1 model while only some evidence was found that CFA was sensitive to sample size and estimator used. The similarities in performance of CFI and TLI found here were not observed in previous simulation studies of MCFA (Navruz, 2016). The distribution of RMSEA was found to be sensitive to sample size, observed variable ICC and estimator. In particular, when the level-1 sample size increased, the values of RMSEA tended to be smaller. When models were estimated with ULSMV or WLSMV, RMSEA values were less variable and lower on average than models estimated with MLR. However, this finding may be due to estimation with MLR assumed the indicators were continuous and not categorical.

The observed estimates of SRMRW and SRMRB were sensitivity to many of the design factors included, such as number of groups, number of units with a group, observed variable ICC, latent variable ICC, and estimator. The effect of each of these design factors is somewhat oversimplified because of the presence of numerous interactions (see Table 4.2). One interaction particular interesting for SRMRW was how the effect of estimator depended on the latent variable ICC. When the latent variable ICC was low, WLSMV and ULSMV resulted in nearly identical distributions of SRMRW, but when the latent variable ICC was high, the SRMRW estimates were higher on average when estimated with ULSMV versus WLSMV. Additionally, these common fit indices are not equally sensitive to misspecification of the level-1 model versus the level-2 model.

The differences in sensitivity to model misspecification at different levels leads to varying degrees of utility of these fit statistics. The difficulty in using these fit statistics for MCFA is that misspecification could be due to the within- or between-group model. This consideration is one of the reason the SRMR statistics are so conceptually useful; that is, they guide researchers on which part of the model may be misspecified. However, I only found the pooled within group statistic (SRMRW) to give useful guidance on potential misspecification of the level-1 model. SRMRB was found to be too variable across conditions and/or models to consistently discriminate between a correctly and incorrectly specified between groups model. The SRMRB statistic also returned some very large values for smaller number of groups. For example, in a condition with a very small number of groups (30) returned an estimated SRMRB for a correctly specified model of 2.54. After examining the output for this model, the only indication that an estimation issue occurred was that several of the items had zero variation across a few groups. This feature of zero variation within group is commonly found in multilevel models especially for groups with few observations and the impact of such circumstances on estimation is unclear (see M*plus* FAQ). Researchers may be able to get fine grained information of the fit of each level's model by using the level specific indices proposed by Ryu and West (2009). However, only one additional study was found that described these level specific statistics (Ryu, 2014). These level-specific indices should be evaluated in future studies for applicability to MCFA and multilevel SEM more generally.

One of the major contributions of this thesis is the use of ROC analysis for investigating the performance of fit statistics for identifying misspecification. ROC analysis provides methodologists a more rigorous tool for examining the results of Monte Carlo simulation studies of model fit statistics for identifying model mispseci-fication. The ROC analysis methods used in this study were fairly basic and could be expanded upon to account for issues such as cell size differences. Using the methods provided will guide researchers into more rigorous examinations of model fit and how fit statistics perform in complex multilevel settings.

### *Recommendations for MCFA Fit Statistics*

The availability of fit indices is intended to aid researchers in diagnosing misfit and ultimately select the correct model for interpretation. Based on the mixed find-ings of the current study, the utility of available fit indices for interpreting MCFA models is may be limited. Researchers are strongly encouraged to seek more ways of evaluating model fit beyond the fit indices commonly reported by M*plus*. The fit statistics described in Ryu and West (2009) appear promising to this end. If one is going to use the commonly reported indices, they should be interpreted with caution, especially when a small number of groups is used. The commonly used cut-off criteria may also be inappropriate depending on the estimator used. The recommended cut-off criteria for CFI and TLI match those of those commonly used when the MCFA model treats the ordered categorical data as continuous and estimated with MLR. When a robust estimator is used different criteria were found to be needed. Across all estimators, I found that SRMRB performed *best* at detecting level-2 misfit but this statistic still does not discriminate between correct and incorrect level-2 models well (see ROC analyses in Chapter 4). A summary of the recommended criteria for the commonly used fit statistics is outlined in Table 5.1.

Table 5.1

*Recommended cut-off criteria for fit indices for categorical data in MCFA*

| Index | MLR | ULSMV | WLSMV |
|---|---|---|---|
| CFI | .95 | .97 | .98 |
| TLI | .95 | .97 | .98 |
| RMSEA | .03 | .02 | .02 |
| SRMRW[1] | .05 | .05 | .05 |
| SRMRB[2] | - | - | - |

*Note.* [1] SRMRW should only be used for finding evidence for the level-1 model specification.
[2] SRMRB is not recommended unless there is strong theoretical justification..

*Delimitations*

As with any simulation study, the results of only generalize to the limited conditions examined. That said, the conditions chosen were selected to mirror conditions of applied researchers as close as possible while still maintaining parsimony. Estimation with MLR resulted in the most usable cases per cell on average, but this may have occurred because data were treated as continuous. Using MLR in M*plus* with categorical data requires numerical integration across four dimensions (one dimension per latent variable) that is computationally burdensome, so I treated these data as continuous in order to estimate these models in a reasonable amount of time. Additionally, the cells of this design ended up with unequal sample sizes due to convergence issues and the number usable replications. In some cells, the number of usable cases was zero (see Table C.2). Proper solutions were checked by looking for negative variances; however, another type of improper solution is impossible estimates of factor loadings. By only checking for negative variances we may be underestimating the rate of admissible cases. These factors that limit the sample size in some cells could have influenced these results in unexpected ways and caused the impact of some design factors to be underestimated. Future work could address how to estimate these types of models under conditions that failed to converge or provide useful information.

### *Future Directions*

The use of only one type of model misspecification (i.e., an omitted crossloading) limits the generalizability of these finding; although, this is a commonly observed type of model misspecification. Other potential sources of misspecification (i.e., residual correlations among items, omitted factor correlation(s), incorrect number of factor(s) specified, etc.) may result in different conclusions about distributions of these fit statistics. Other types of misspecification can be investigated with these data, and the generated data are available online (Padgett, 2019).

In this study, scope was limited to fit the performance of common fit statistics to detect misspecification. Although, as already mentioned, there is a growing interest in level specific fit indices for identifying more nuanced information on misfit (Ryu & West, 2009). Future investigations of fit statistics in MCFA and multilevel SEM would benefit from a deeper investigation into the performance of these statistics. The estimation of these statistics are not easily available for investigation so future work to clearly outline the estimation (e.g., provide code) would likely be of interest to a wide range of researchers.

The use of level specific fit measures may help researchers find evidence for their model when the measurement model is not invariant across levels. In this study, only one type of multilevel measurement model was investigated where the level-2 model reflects an aggregate of the level-1 measurement model. However, this may not always be the case and allowing for varying covariance structures across levels may also influence the performance of fit statistics since the level-2 model isn't constrained to have the same loadings as the level-1 model. Modeling different structure across levels is becoming more common in applied examples of MCFA/MSEM when the research interest is in level-2 constructs (Kim, Dedrick, Cao, & Ferron, 2016). Future research should investigate methods for identifying when a more general structure is necessary across levels. This line of research would need to identify considerations from testing measurement invariance across levels and how to identify when a different structure at level-2 if necessary.

APPENDICES

# APPENDIX A

## R Code for Threshold Generation

The following was used to generate threshold values for generating an observed variance of 1 based on categorical responses. The code is generalized such that the total variance is calculated for any number of cross loadings. The user needs to make sure to set up each piece correctly with the information corresponding to population model.

```
## Set up of population model specification
lambda_w <- matrix(c(.6,.6), ncol=2) #  factor loading(s) of
   level-1
phi_w <- matrix(c(1,.3,.3,1),ncol=2) # factor variances level
   -1
psi_w <- matrix(c(1)) # residual variances level-1
lambda_b <- matrix(c(.6,.6), ncol=2) #  factor loading(s) of
   level-2
phi_b <- matrix(c(.25,.075,.075,.25),ncol=2) # factor
   variances level-2
psi_b <- matrix(c(.25)) # residual variances level-2
## Total variance of latent response for a cross loaded item
## Note: %*% is matrix multiplication
Vyi <- lambda_w%*%phi_w%*%t(lambda_w) + psi_w + lambda_b%*%phi
   _b%*%t(lambda_b) + psi_b
## Vector of Coded values for each category
## Note: coded values are arbitrary as long as sequential
X <- 1:5
## Vector of Response Probabilities
P <- c(.0675, .2325, .4, .2325, .0675)
## Mean
sum(P*X)
## Variance of Observed Categorical Items
sum(P*X**2) - (sum(P*X))**2
## Threshold Generation
t <-numeric(4)
i <- 1
while(i <= length(P) - 1){
  t[i] <- qnorm(sum(P[1:i]), 0, sqrt(Vyi))
  i <- i + 1
}
t
```

# APPENDIX B

## Additional effect size estimates for ANOVAs

Table B.1

*Summary of ANOVA by effect size estimates with $\omega^2$*

| Effect | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|
| $N_1$ | 0.040 | 0.040 | **0.051** | **0.187** | **0.052** |
| $N_2$ | 0.013 | 0.013 | 0.006 | **0.193** | **0.329** |
| $ICC_O$ | 0.005 | 0.005 | 0.042 | 0.018 | **0.136** |
| $ICC_L$ | 0.004 | 0.004 | 0.002 | 0.004 | 0.038 |
| Model | **0.317** | **0.315** | **0.390** | **0.324** | 0.020 |
| Estimator | 0.024 | 0.024 | **0.092** | **0.079** | **0.056** |
| $N_1$:$N_2$ | 0.036 | 0.036 | 0.017 | 0.038 | 0.006 |
| $N_1$:$ICC_O$ | 0.005 | 0.005 | 0.001 | 0.000 | 0.039 |
| $N_1$:$ICC_L$ | 0.000 | 0.000 | 0.002 | 0.000 | 0.011 |
| $N_1$:Model | 0.001 | 0.001 | 0.003 | 0.012 | 0.000 |
| $N_1$:Estimator | 0.002 | 0.002 | 0.001 | 0.002 | 0.000 |
| $N_2$:$ICC_O$ | 0.001 | 0.001 | 0.005 | 0.001 | 0.004 |
| $N_2$:$ICC_L$ | 0.005 | 0.006 | 0.001 | 0.001 | 0.015 |
| $N_2$:Model | 0.003 | 0.003 | 0.012 | 0.010 | 0.001 |
| $N_2$:Estimator | 0.034 | 0.034 | 0.046 | 0.008 | 0.002 |
| $ICC_O$:$ICC_L$ | 0.003 | 0.003 | 0.002 | 0.000 | 0.032 |
| $ICC_O$:Model | 0.007 | 0.007 | 0.017 | 0.001 | 0.006 |
| $ICC_O$:Estimator | 0.006 | 0.006 | 0.011 | 0.009 | 0.000 |
| $ICC_L$:Model | 0.018 | 0.018 | 0.019 | 0.002 | 0.012 |
| $ICC_L$:Estimator | 0.013 | 0.013 | 0.001 | 0.012 | 0.002 |
| Model:Estimator | 0.020 | 0.020 | 0.018 | 0.003 | 0.001 |

*Note.* All values greater than .05 (i.e., more than 5% of variability explained) were bolded. The meaning of each value can be interpreted as follows for the $N_1$ (number of level-1 units) effect. For CFI, 4% of the variability in observed scores can be attributed to the number of level-1 units ($N_1$) were sampled per group.

# APPENDIX C

## Convergence and Admissibility Across Conditions

### *Convergence Breakdown*

This sections contains the breakdown of convergence rates across all conditions, models, and estimators.

Table C.1

*Convergence Across All Conditions, Models, and Estimators*

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| C | MLR | 5 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 5 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 5 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| C | MLR | 10 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 10 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 10 | 200 | 1.000 | 0.998 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 30 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | MLR | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 1.000 |
| C | MLR | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 1.000 |
| C | MLR | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 1.000 |
| C | ULSMV | 5 | 30 | 0.998 | 0.998 | 0.966 | 1.000 | 1.000 | 0.976 |
| C | ULSMV | 5 | 50 | 1.000 | 1.000 | 0.998 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 5 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 10 | 30 | 1.000 | 1.000 | 0.998 | 1.000 | 1.000 | 0.996 |
| C | ULSMV | 10 | 50 | 1.000 | 1.000 | 1.000 | 0.998 | 1.000 | 1.000 |
| C | ULSMV | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 30 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| C | ULSMV | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 5 | 30 | 0.992 | 1.000 | 0.998 | 1.000 | 1.000 | 0.996 |
| C | WLSMV | 5 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 5 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

(continued)

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| C | WLSMV | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 10 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 10 | 50 | 1.000 | 1.000 | 1.000 | 0.998 | 1.000 | 1.000 |
| C | WLSMV | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 30 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M1 | MLR | 5 | 30 | 0.980 | 0.972 | 0.956 | 1.000 | 0.996 | 0.988 |
| M1 | MLR | 5 | 50 | 0.992 | 0.972 | 0.958 | 0.998 | 1.000 | 0.996 |
| M1 | MLR | 5 | 100 | 0.982 | 0.972 | 0.942 | 1.000 | 1.000 | 0.996 |
| M1 | MLR | 5 | 200 | 0.996 | 0.978 | 0.942 | 1.000 | 1.000 | 1.000 |
| M1 | MLR | 10 | 30 | 0.992 | 0.974 | 0.932 | 1.000 | 1.000 | 0.992 |
| M1 | MLR | 10 | 50 | 0.990 | 0.954 | 0.926 | 1.000 | 1.000 | 1.000 |
| M1 | MLR | 10 | 100 | 0.992 | 0.964 | 0.924 | 1.000 | 1.000 | 1.000 |
| M1 | MLR | 10 | 200 | 0.992 | 0.992 | 0.940 | 1.000 | 1.000 | 1.000 |
| M1 | MLR | 30 | 30 | 0.984 | 0.956 | 0.894 | 1.000 | 1.000 | 0.994 |
| M1 | MLR | 30 | 50 | 0.990 | 0.966 | 0.908 | 1.000 | 1.000 | 0.996 |
| M1 | MLR | 30 | 100 | 0.994 | 0.986 | 0.928 | 1.000 | 1.000 | 1.000 |
| M1 | MLR | 30 | 200 | 1.000 | 0.994 | 0.932 | 1.000 | 0.998 | 1.000 |
| M1 | ULSMV | 5 | 30 | 0.924 | 0.892 | 0.866 | 0.980 | 0.970 | 0.938 |
| M1 | ULSMV | 5 | 50 | 0.932 | 0.916 | 0.916 | 0.998 | 0.996 | 0.988 |
| M1 | ULSMV | 5 | 100 | 0.966 | 0.942 | 0.938 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 5 | 200 | 0.986 | 0.960 | 0.962 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 10 | 30 | 0.948 | 0.898 | 0.938 | 0.998 | 0.996 | 0.984 |
| M1 | ULSMV | 10 | 50 | 0.956 | 0.936 | 0.932 | 0.998 | 1.000 | 0.990 |
| M1 | ULSMV | 10 | 100 | 0.980 | 0.958 | 0.958 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 10 | 200 | 1.000 | 0.976 | 0.952 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 30 | 30 | 0.982 | 0.940 | 0.932 | 1.000 | 1.000 | 0.990 |
| M1 | ULSMV | 30 | 50 | 0.992 | 0.964 | 0.948 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 30 | 100 | 1.000 | 0.984 | 0.954 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 30 | 200 | 1.000 | 0.986 | 0.972 | 0.998 | 1.000 | 1.000 |
| M1 | WLSMV | 5 | 30 | 0.866 | 0.842 | 0.866 | 0.968 | 0.960 | 0.920 |
| M1 | WLSMV | 5 | 50 | 0.892 | 0.904 | 0.898 | 0.996 | 0.990 | 0.968 |
| M1 | WLSMV | 5 | 100 | 0.954 | 0.946 | 0.932 | 1.000 | 1.000 | 0.992 |
| M1 | WLSMV | 5 | 200 | 0.984 | 0.954 | 0.946 | 1.000 | 1.000 | 1.000 |
| M1 | WLSMV | 10 | 30 | 0.920 | 0.900 | 0.880 | 0.994 | 0.996 | 0.978 |
| M1 | WLSMV | 10 | 50 | 0.964 | 0.932 | 0.884 | 0.998 | 1.000 | 0.998 |
| M1 | WLSMV | 10 | 100 | 0.984 | 0.946 | 0.918 | 1.000 | 1.000 | 1.000 |
| M1 | WLSMV | 10 | 200 | 0.996 | 0.980 | 0.940 | 1.000 | 1.000 | 1.000 |

(continued)

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| M1 | WLSMV | 30 | 30 | 0.982 | 0.934 | 0.902 | 1.000 | 0.994 | 0.980 |
| M1 | WLSMV | 30 | 50 | 0.992 | 0.946 | 0.914 | 1.000 | 1.000 | 0.998 |
| M1 | WLSMV | 30 | 100 | 0.998 | 0.974 | 0.936 | 1.000 | 1.000 | 1.000 |
| M1 | WLSMV | 30 | 200 | 1.000 | 0.982 | 0.926 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 5 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| M2 | MLR | 5 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 5 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 10 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 10 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| M2 | MLR | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 30 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 1.000 |
| M2 | MLR | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 | 1.000 |
| M2 | MLR | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 1.000 |
| M2 | ULSMV | 5 | 30 | 0.998 | 0.994 | 0.968 | 0.938 | 0.934 | 0.914 |
| M2 | ULSMV | 5 | 50 | 1.000 | 1.000 | 1.000 | 0.982 | 0.966 | 0.944 |
| M2 | ULSMV | 5 | 100 | 1.000 | 1.000 | 1.000 | 0.998 | 0.988 | 0.970 |
| M2 | ULSMV | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 |
| M2 | ULSMV | 10 | 30 | 1.000 | 1.000 | 0.990 | 0.980 | 0.940 | 0.944 |
| M2 | ULSMV | 10 | 50 | 1.000 | 1.000 | 1.000 | 0.992 | 0.982 | 0.956 |
| M2 | ULSMV | 10 | 100 | 1.000 | 1.000 | 1.000 | 0.998 | 0.998 | 0.978 |
| M2 | ULSMV | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| M2 | ULSMV | 30 | 30 | 1.000 | 1.000 | 1.000 | 0.988 | 0.962 | 0.934 |
| M2 | ULSMV | 30 | 50 | 1.000 | 1.000 | 1.000 | 0.992 | 0.988 | 0.972 |
| M2 | ULSMV | 30 | 100 | 1.000 | 1.000 | 1.000 | 0.998 | 0.992 | 0.982 |
| M2 | ULSMV | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 5 | 30 | 0.994 | 0.996 | 0.998 | 0.998 | 1.000 | 0.998 |
| M2 | WLSMV | 5 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 5 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 10 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 10 | 50 | 1.000 | 1.000 | 1.000 | 0.998 | 1.000 | 1.000 |
| M2 | WLSMV | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 5 | 30 | 1.000 | 1.000 | 0.998 | 1.000 | 1.000 | 1.000 |

(continued)

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| M12 | MLR | 5 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 5 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 10 | 30 | 1.000 | 0.998 | 1.000 | 1.000 | 1.000 | 0.998 |
| M12 | MLR | 10 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 30 | 30 | 1.000 | 1.000 | 1.000 | 1.000 | 0.994 | 1.000 |
| M12 | MLR | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 | 1.000 |
| M12 | MLR | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | ULSMV | 5 | 30 | 0.976 | 0.952 | 0.946 | 0.942 | 0.940 | 0.906 |
| M12 | ULSMV | 5 | 50 | 0.988 | 0.988 | 0.978 | 0.974 | 0.964 | 0.942 |
| M12 | ULSMV | 5 | 100 | 0.998 | 1.000 | 0.998 | 0.996 | 0.988 | 0.958 |
| M12 | ULSMV | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.996 |
| M12 | ULSMV | 10 | 30 | 0.990 | 0.990 | 0.980 | 0.970 | 0.952 | 0.938 |
| M12 | ULSMV | 10 | 50 | 0.998 | 1.000 | 1.000 | 0.994 | 0.986 | 0.952 |
| M12 | ULSMV | 10 | 100 | 1.000 | 1.000 | 1.000 | 0.998 | 0.998 | 0.972 |
| M12 | ULSMV | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| M12 | ULSMV | 30 | 30 | 1.000 | 1.000 | 1.000 | 0.992 | 0.968 | 0.924 |
| M12 | ULSMV | 30 | 50 | 1.000 | 1.000 | 1.000 | 0.992 | 0.982 | 0.962 |
| M12 | ULSMV | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 0.992 | 0.986 |
| M12 | ULSMV | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.998 |
| M12 | WLSMV | 5 | 30 | 0.926 | 0.938 | 0.942 | 0.976 | 0.972 | 0.954 |
| M12 | WLSMV | 5 | 50 | 0.964 | 0.974 | 0.976 | 0.994 | 0.980 | 0.984 |
| M12 | WLSMV | 5 | 100 | 0.996 | 1.000 | 1.000 | 1.000 | 1.000 | 0.994 |
| M12 | WLSMV | 5 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | WLSMV | 10 | 30 | 0.982 | 0.986 | 0.964 | 0.994 | 0.992 | 0.980 |
| M12 | WLSMV | 10 | 50 | 0.998 | 1.000 | 0.998 | 0.998 | 1.000 | 0.996 |
| M12 | WLSMV | 10 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | WLSMV | 10 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | WLSMV | 30 | 30 | 1.000 | 0.998 | 0.982 | 1.000 | 0.994 | 0.988 |
| M12 | WLSMV | 30 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | WLSMV | 30 | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| M12 | WLSMV | 30 | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

### *Admissible Solutions Breakdown*

Similar to the previous section, this section contains the breakdown for admissible solutions across all conditions, models, and estimators.

Table C.2

*Admissibility Across All Conditions, Models, and Estimators*

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| C | MLR | 5 | 30 | 0.522 | 0.468 | 0.360 | 0.372 | 0.792 | 0.938 |
| C | MLR | 5 | 50 | 0.560 | 0.548 | 0.488 | 0.500 | 0.972 | 0.996 |
| C | MLR | 5 | 100 | 0.752 | 0.760 | 0.600 | 0.726 | 0.996 | 1.000 |
| C | MLR | 5 | 200 | 0.952 | 0.890 | 0.744 | 0.956 | 1.000 | 1.000 |
| C | MLR | 10 | 30 | 0.608 | 0.584 | 0.446 | 0.494 | 0.988 | 0.988 |
| C | MLR | 10 | 50 | 0.832 | 0.698 | 0.482 | 0.762 | 1.000 | 1.000 |
| C | MLR | 10 | 100 | 0.970 | 0.890 | 0.672 | 0.988 | 1.000 | 1.000 |
| C | MLR | 10 | 200 | 1.000 | 0.984 | 0.796 | 1.000 | 1.000 | 1.000 |
| C | MLR | 30 | 30 | 0.950 | 0.688 | 0.460 | 0.938 | 1.000 | 0.994 |
| C | MLR | 30 | 50 | 0.996 | 0.830 | 0.592 | 0.994 | 1.000 | 1.000 |
| C | MLR | 30 | 100 | 1.000 | 0.962 | 0.716 | 1.000 | 1.000 | 1.000 |
| C | MLR | 30 | 200 | 1.000 | 0.998 | 0.858 | 1.000 | 1.000 | 1.000 |
| C | ULSMV | 5 | 30 | **.030** | 0.357 | 0.335 | 0.036 | 0.730 | 0.953 |
| C | ULSMV | 5 | 50 | 0.138 | 0.542 | 0.483 | 0.092 | 0.938 | 0.998 |
| C | ULSMV | 5 | 100 | 0.494 | 0.758 | 0.596 | 0.314 | 0.990 | 1.000 |
| C | ULSMV | 5 | 200 | 0.890 | 0.880 | 0.754 | 0.728 | 1.000 | 1.000 |
| C | ULSMV | 10 | 30 | 0.352 | 0.574 | 0.449 | 0.134 | 0.952 | 0.982 |
| C | ULSMV | 10 | 50 | 0.668 | 0.698 | 0.498 | 0.387 | 0.994 | 1.000 |
| C | ULSMV | 10 | 100 | 0.932 | 0.884 | 0.668 | 0.754 | 1.000 | 1.000 |
| C | ULSMV | 10 | 200 | 0.998 | 0.980 | 0.790 | 0.978 | 1.000 | 1.000 |
| C | ULSMV | 30 | 30 | 0.902 | 0.690 | 0.482 | 0.518 | 0.990 | 0.992 |
| C | ULSMV | 30 | 50 | 0.988 | 0.848 | 0.598 | 0.844 | 1.000 | 1.000 |
| C | ULSMV | 30 | 100 | 1.000 | 0.960 | 0.738 | 0.970 | 1.000 | 1.000 |
| C | ULSMV | 30 | 200 | 1.000 | 1.000 | 0.846 | 1.000 | 1.000 | 1.000 |
| C | WLSMV | 5 | 30 | 0.048 | 0.250 | 0.224 | 0.024 | 0.674 | 0.914 |
| C | WLSMV | 5 | 50 | 0.146 | 0.404 | 0.358 | 0.052 | 0.868 | 0.992 |
| C | WLSMV | 5 | 100 | 0.556 | 0.682 | 0.532 | 0.238 | 0.982 | 0.998 |
| C | WLSMV | 5 | 200 | 0.900 | 0.862 | 0.716 | 0.614 | 1.000 | 1.000 |
| C | WLSMV | 10 | 30 | 0.364 | 0.418 | 0.358 | 0.080 | 0.854 | 0.974 |
| C | WLSMV | 10 | 50 | 0.726 | 0.618 | 0.424 | 0.218 | 0.976 | 1.000 |
| C | WLSMV | 10 | 100 | 0.946 | 0.846 | 0.638 | 0.518 | 0.998 | 1.000 |
| C | WLSMV | 10 | 200 | 1.000 | 0.980 | 0.778 | 0.856 | 1.000 | 1.000 |
| C | WLSMV | 30 | 30 | 0.890 | 0.636 | 0.432 | 0.234 | 0.950 | 0.988 |
| C | WLSMV | 30 | 50 | 0.990 | 0.816 | 0.560 | 0.552 | 1.000 | 1.000 |
| C | WLSMV | 30 | 100 | 1.000 | 0.962 | 0.692 | 0.812 | 1.000 | 1.000 |
| C | WLSMV | 30 | 200 | 1.000 | 0.998 | 0.860 | 0.982 | 1.000 | 1.000 |
| M1 | MLR | 5 | 30 | 0.327 | 0.401 | 0.400 | 0.266 | 0.629 | 0.783 |

(continued)

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| M1 | MLR | 5 | 50 | 0.302 | 0.508 | 0.524 | 0.321 | 0.818 | 0.916 |
| M1 | MLR | 5 | 100 | 0.271 | 0.611 | 0.599 | 0.274 | 0.906 | 0.992 |
| M1 | MLR | 5 | 200 | 0.243 | 0.714 | 0.758 | 0.250 | 0.990 | 1.000 |
| M1 | MLR | 10 | 30 | 0.385 | 0.556 | 0.474 | 0.364 | 0.878 | 0.895 |
| M1 | MLR | 10 | 50 | 0.453 | 0.665 | 0.551 | 0.514 | 0.940 | 0.968 |
| M1 | MLR | 10 | 100 | 0.556 | 0.790 | 0.680 | 0.762 | 0.990 | 0.998 |
| M1 | MLR | 10 | 200 | 0.639 | 0.895 | 0.787 | 0.896 | 1.000 | 1.000 |
| M1 | MLR | 30 | 30 | 0.764 | 0.646 | 0.483 | 0.852 | 0.934 | 0.905 |
| M1 | MLR | 30 | 50 | 0.855 | 0.764 | 0.586 | 0.976 | 0.984 | 0.980 |
| M1 | MLR | 30 | 100 | 0.958 | 0.903 | 0.733 | 0.998 | 1.000 | 1.000 |
| M1 | MLR | 30 | 200 | 0.990 | 0.948 | 0.800 | 1.000 | 1.000 | 1.000 |
| M1 | ULSMV | 5 | 30 | 0.041 | 0.377 | 0.374 | 0.012 | 0.542 | 0.823 |
| M1 | ULSMV | 5 | 50 | 0.116 | 0.533 | 0.537 | 0.022 | 0.661 | 0.925 |
| M1 | ULSMV | 5 | 100 | 0.393 | 0.709 | 0.623 | 0.026 | 0.708 | 0.988 |
| M1 | ULSMV | 5 | 200 | 0.684 | 0.871 | 0.728 | 0.002 | 0.822 | 1.000 |
| M1 | ULSMV | 10 | 30 | 0.325 | 0.570 | 0.507 | 0.018 | 0.649 | 0.904 |
| M1 | ULSMV | 10 | 50 | 0.494 | 0.697 | 0.552 | 0.020 | 0.658 | 0.964 |
| M1 | ULSMV | 10 | 100 | 0.737 | 0.837 | 0.678 | 0.002 | 0.770 | 0.994 |
| M1 | ULSMV | 10 | 200 | 0.834 | 0.947 | 0.796 | **0.000** | 0.858 | 1.000 |
| M1 | ULSMV | 30 | 30 | 0.697 | 0.672 | 0.513 | 0.008 | 0.666 | 0.915 |
| M1 | ULSMV | 30 | 50 | 0.792 | 0.805 | 0.633 | 0.002 | 0.794 | 0.976 |
| M1 | ULSMV | 30 | 100 | 0.878 | 0.933 | 0.757 | 0.000 | 0.834 | 1.000 |
| M1 | ULSMV | 30 | 200 | 0.944 | 0.976 | 0.800 | 0.000 | 0.916 | 1.000 |
| M1 | WLSMV | 5 | 30 | 0.046 | 0.271 | 0.293 | 0.008 | 0.471 | 0.785 |
| M1 | WLSMV | 5 | 50 | 0.141 | 0.423 | 0.439 | 0.006 | 0.495 | 0.826 |
| M1 | WLSMV | 5 | 100 | 0.421 | 0.658 | 0.575 | 0.010 | 0.462 | 0.933 |
| M1 | WLSMV | 5 | 200 | 0.720 | 0.855 | 0.698 | 0.000 | 0.444 | 0.962 |
| M1 | WLSMV | 10 | 30 | 0.313 | 0.436 | 0.405 | 0.010 | 0.494 | 0.816 |
| M1 | WLSMV | 10 | 50 | 0.581 | 0.622 | 0.489 | 0.006 | 0.438 | 0.892 |
| M1 | WLSMV | 10 | 100 | 0.752 | 0.801 | 0.627 | 0.000 | 0.348 | 0.914 |
| M1 | WLSMV | 10 | 200 | 0.878 | 0.939 | 0.787 | **.000** | 0.280 | 0.992 |
| M1 | WLSMV | 30 | 30 | 0.690 | 0.580 | 0.446 | 0.004 | 0.400 | 0.718 |
| M1 | WLSMV | 30 | 50 | 0.835 | 0.784 | 0.593 | 0.000 | 0.370 | 0.842 |
| M1 | WLSMV | 30 | 100 | 0.900 | 0.916 | 0.718 | 0.000 | 0.258 | 0.946 |
| M1 | WLSMV | 30 | 200 | 0.962 | 0.984 | 0.801 | 0.000 | 0.166 | 0.992 |
| M2 | MLR | 5 | 30 | 0.496 | 0.438 | 0.354 | 0.382 | 0.822 | 0.946 |
| M2 | MLR | 5 | 50 | 0.562 | 0.538 | 0.476 | 0.550 | 0.978 | 0.996 |
| M2 | MLR | 5 | 100 | 0.764 | 0.702 | 0.576 | 0.772 | 1.000 | 1.000 |
| M2 | MLR | 5 | 200 | 0.926 | 0.860 | 0.732 | 0.974 | 1.000 | 1.000 |
| M2 | MLR | 10 | 30 | 0.580 | 0.574 | 0.456 | 0.530 | 0.986 | 0.982 |
| M2 | MLR | 10 | 50 | 0.802 | 0.694 | 0.480 | 0.806 | 1.000 | 1.000 |

(continued)

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| M2 | MLR | 10 | 100 | 0.966 | 0.858 | 0.642 | 0.990 | 1.000 | 1.000 |
| M2 | MLR | 10 | 200 | 1.000 | 0.972 | 0.786 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 30 | 30 | 0.936 | 0.674 | 0.484 | 0.952 | 1.000 | 0.998 |
| M2 | MLR | 30 | 50 | 0.996 | 0.826 | 0.562 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 30 | 100 | 1.000 | 0.962 | 0.718 | 1.000 | 1.000 | 1.000 |
| M2 | MLR | 30 | 200 | 1.000 | 0.992 | 0.832 | 1.000 | 1.000 | 1.000 |
| M2 | ULSMV | 5 | 30 | 0.046 | 0.368 | 0.341 | 0.051 | 0.709 | 0.921 |
| M2 | ULSMV | 5 | 50 | 0.154 | 0.550 | 0.486 | 0.145 | 0.919 | 0.985 |
| M2 | ULSMV | 5 | 100 | 0.514 | 0.736 | 0.582 | 0.427 | 0.988 | 0.996 |
| M2 | ULSMV | 5 | 200 | 0.920 | 0.888 | 0.756 | 0.820 | 0.998 | 1.000 |
| M2 | ULSMV | 10 | 30 | 0.372 | 0.584 | 0.475 | 0.206 | 0.953 | 0.968 |
| M2 | ULSMV | 10 | 50 | 0.696 | 0.688 | 0.494 | 0.472 | 0.978 | 0.990 |
| M2 | ULSMV | 10 | 100 | 0.954 | 0.886 | 0.654 | 0.834 | 0.996 | 0.998 |
| M2 | ULSMV | 10 | 200 | 0.998 | 0.976 | 0.798 | 0.982 | 1.000 | 1.000 |
| M2 | ULSMV | 30 | 30 | 0.922 | 0.676 | 0.484 | 0.605 | 0.967 | 0.989 |
| M2 | ULSMV | 30 | 50 | 0.994 | 0.828 | 0.566 | 0.857 | 0.992 | 0.998 |
| M2 | ULSMV | 30 | 100 | 1.000 | 0.966 | 0.718 | 0.966 | 1.000 | 1.000 |
| M2 | ULSMV | 30 | 200 | 1.000 | 0.992 | 0.832 | 1.000 | 1.000 | 1.000 |
| M2 | WLSMV | 5 | 30 | 0.058 | 0.237 | 0.210 | 0.044 | 0.694 | 0.918 |
| M2 | WLSMV | 5 | 50 | 0.170 | 0.424 | 0.392 | 0.094 | 0.910 | 0.996 |
| M2 | WLSMV | 5 | 100 | 0.550 | 0.682 | 0.532 | 0.328 | 0.996 | 1.000 |
| M2 | WLSMV | 5 | 200 | 0.920 | 0.850 | 0.720 | 0.716 | 1.000 | 1.000 |
| M2 | WLSMV | 10 | 30 | 0.378 | 0.410 | 0.390 | 0.118 | 0.908 | 0.974 |
| M2 | WLSMV | 10 | 50 | 0.726 | 0.614 | 0.438 | 0.273 | 0.990 | 1.000 |
| M2 | WLSMV | 10 | 100 | 0.966 | 0.846 | 0.614 | 0.614 | 1.000 | 1.000 |
| M2 | WLSMV | 10 | 200 | 1.000 | 0.974 | 0.780 | 0.922 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 30 | 0.914 | 0.638 | 0.434 | 0.262 | 0.970 | 0.990 |
| M2 | WLSMV | 30 | 50 | 0.994 | 0.820 | 0.550 | 0.548 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 100 | 1.000 | 0.956 | 0.692 | 0.838 | 1.000 | 1.000 |
| M2 | WLSMV | 30 | 200 | 1.000 | 0.994 | 0.828 | 0.996 | 1.000 | 1.000 |
| M12 | MLR | 5 | 30 | 0.510 | 0.410 | 0.345 | 0.338 | 0.764 | 0.918 |
| M12 | MLR | 5 | 50 | 0.552 | 0.548 | 0.492 | 0.494 | 0.930 | 0.986 |
| M12 | MLR | 5 | 100 | 0.752 | 0.746 | 0.608 | 0.776 | 0.988 | 0.998 |
| M12 | MLR | 5 | 200 | 0.936 | 0.886 | 0.754 | 0.982 | 1.000 | 1.000 |
| M12 | MLR | 10 | 30 | 0.602 | 0.591 | 0.432 | 0.528 | 0.946 | 0.964 |
| M12 | MLR | 10 | 50 | 0.828 | 0.700 | 0.486 | 0.786 | 0.988 | 0.998 |
| M12 | MLR | 10 | 100 | 0.972 | 0.880 | 0.654 | 0.992 | 1.000 | 1.000 |
| M12 | MLR | 10 | 200 | 0.998 | 0.980 | 0.786 | 1.000 | 1.000 | 1.000 |
| M12 | MLR | 30 | 30 | 0.944 | 0.694 | 0.498 | 0.938 | 0.984 | 0.982 |
| M12 | MLR | 30 | 50 | 0.992 | 0.836 | 0.554 | 1.000 | 0.996 | 1.000 |
| M12 | MLR | 30 | 100 | 1.000 | 0.970 | 0.714 | 1.000 | 1.000 | 1.000 |

(continued)

| | | | | $ICC_L = .1$ | | | $ICC_L = .5$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Estimator | $N_1$ | $N_2$ | .1 | .3 | .5 | .1 | .3 | .5 |
| M12 | MLR | 30 | 200 | 1.000 | 0.994 | 0.836 | 1.000 | 1.000 | 1.000 |
| M12 | ULSMV | 5 | 30 | 0.041 | 0.363 | 0.323 | 0.023 | 0.621 | 0.883 |
| M12 | ULSMV | 5 | 50 | 0.150 | 0.498 | 0.497 | 0.066 | 0.751 | 0.947 |
| M12 | ULSMV | 5 | 100 | 0.453 | 0.730 | 0.599 | 0.066 | 0.796 | 0.996 |
| M12 | ULSMV | 5 | 200 | 0.820 | 0.880 | 0.756 | 0.068 | 0.864 | 1.000 |
| M12 | ULSMV | 10 | 30 | 0.352 | 0.560 | 0.435 | 0.062 | 0.742 | 0.940 |
| M12 | ULSMV | 10 | 50 | 0.627 | 0.702 | 0.482 | 0.085 | 0.769 | 0.975 |
| M12 | ULSMV | 10 | 100 | 0.844 | 0.888 | 0.658 | 0.052 | 0.840 | 0.994 |
| M12 | ULSMV | 10 | 200 | 0.938 | 0.976 | 0.794 | 0.016 | 0.904 | 1.000 |
| M12 | ULSMV | 30 | 30 | 0.780 | 0.686 | 0.498 | 0.085 | 0.787 | 0.961 |
| M12 | ULSMV | 30 | 50 | 0.906 | 0.838 | 0.566 | 0.056 | 0.841 | 0.994 |
| M12 | ULSMV | 30 | 100 | 0.970 | 0.970 | 0.716 | 0.010 | 0.901 | 1.000 |
| M12 | ULSMV | 30 | 200 | 0.998 | 0.992 | 0.832 | 0.000 | 0.954 | 1.000 |
| M12 | WLSMV | 5 | 30 | 0.067 | 0.249 | 0.202 | 0.039 | 0.588 | 0.853 |
| M12 | WLSMV | 5 | 50 | 0.162 | 0.411 | 0.367 | 0.072 | 0.649 | 0.911 |
| M12 | WLSMV | 5 | 100 | 0.462 | 0.666 | 0.532 | 0.108 | 0.708 | 0.968 |
| M12 | WLSMV | 5 | 200 | 0.836 | 0.840 | 0.706 | 0.138 | 0.778 | 0.996 |
| M12 | WLSMV | 10 | 30 | 0.371 | 0.391 | 0.384 | 0.064 | 0.690 | 0.898 |
| M12 | WLSMV | 10 | 50 | 0.651 | 0.622 | 0.435 | 0.120 | 0.712 | 0.956 |
| M12 | WLSMV | 10 | 100 | 0.888 | 0.836 | 0.610 | 0.134 | 0.742 | 0.972 |
| M12 | WLSMV | 10 | 200 | 0.962 | 0.970 | 0.774 | 0.082 | 0.802 | 1.000 |
| M12 | WLSMV | 30 | 30 | 0.842 | 0.639 | 0.430 | 0.138 | 0.686 | 0.874 |
| M12 | WLSMV | 30 | 50 | 0.950 | 0.806 | 0.530 | 0.162 | 0.740 | 0.932 |
| M12 | WLSMV | 30 | 100 | 0.990 | 0.964 | 0.686 | 0.110 | 0.812 | 0.990 |
| M12 | WLSMV | 30 | 200 | 0.998 | 0.996 | 0.818 | 0.036 | 0.826 | 1.000 |

# APPENDIX D

## Fit Statistic Summary Across Conditions

In this Appendix, the fit statistics are broken down to the lowest level units (i.e., each individual condition, model, and estimator). This creates a large number of values that needed to be summarized. In total, there are 72 conditions $\times$ 4 models $\times$ 3 estimators $\times$ 500 replications for a total of 432,000 potential values per fit statistic. The actual number depends on convergence and admissibility. Below are the breakdown where each table is parsed out by model, estimator, latent variable ICC ($\text{ICC}_L$), and observed variable ICC ($\text{ICC}_O$).

Table D.1

*Summary of Fit Statistics Across Conditions: Model C, Estimator MLR, $\text{ICC}_O$ 0.1 and $\text{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.35 | 0.88(0.09) | 0.85(0.11) | 0.05(0.02) | 0.06(0.01) | 0.32(0.05) |
| 50 | 5 | 0.53 | 0.94(0.04) | 0.93(0.05) | 0.03(0.02) | 0.05(0.01) | 0.27(0.04) |
| 100 | 5 | 0.7 | 0.98(0.02) | 0.97(0.02) | 0.02(0.01) | 0.03(0) | 0.22(0.03) |
| 200 | 5 | 0.9 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.17(0.02) |
| 30 | 10 | 0.44 | 0.94(0.04) | 0.93(0.04) | 0.03(0.01) | 0.04(0.01) | 0.25(0.04) |
| 50 | 10 | 0.68 | 0.97(0.02) | 0.97(0.03) | 0.02(0.01) | 0.03(0) | 0.21(0.03) |
| 100 | 10 | 0.88 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.15(0.02) |
| 200 | 10 | 0.91 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.11(0.02) |
| 30 | 30 | 0.61 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.19(0.03) |
| 50 | 30 | 0.84 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.14(0.02) |
| 100 | 30 | 0.91 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.1(0.01) |
| 200 | 30 | 0.94 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.07(0.01) |

Table D.2

*Summary of Fit Statistics Across Conditions: Model C, Estimator MLR, $ICC_O$ 0.1 and $ICC_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.44 | 0.93(0.04) | 0.92(0.05) | 0.04(0.02) | 0.07(0.01) | 0.16(0.04) |
| 50 | 5 | 0.69 | 0.97(0.03) | 0.97(0.03) | 0.03(0.02) | 0.05(0.01) | 0.12(0.02) |
| 100 | 5 | 0.84 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.09(0.02) |
| 200 | 5 | 0.9 | 1(0) | 1(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 0.49 | 0.96(0.03) | 0.95(0.03) | 0.03(0.01) | 0.04(0.01) | 0.12(0.03) |
| 50 | 10 | 0.71 | 0.98(0.02) | 0.98(0.02) | 0.02(0.01) | 0.03(0) | 0.1(0.02) |
| 100 | 10 | 0.89 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.07(0.01) |
| 200 | 10 | 0.92 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 0.62 | 0.99(0.01) | 0.98(0.01) | 0.01(0.01) | 0.02(0) | 0.1(0.02) |
| 50 | 30 | 0.85 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.07(0.01) |
| 100 | 30 | 0.9 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |
| 200 | 30 | 0.93 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.04(0.01) |

Table D.3

*Summary of Fit Statistics Across Conditions: Model C, Estimator MLR, $ICC_O$ 0.3 and $ICC_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.5 | 0.88(0.09) | 0.85(0.11) | 0.04(0.02) | 0.06(0.01) | 0.23(0.03) |
| 50 | 5 | 0.75 | 0.95(0.05) | 0.94(0.06) | 0.02(0.02) | 0.05(0.01) | 0.19(0.02) |
| 100 | 5 | 0.91 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.03(0) | 0.14(0.02) |
| 200 | 5 | 0.9 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.09(0.01) |
| 30 | 10 | 0.63 | 0.95(0.04) | 0.93(0.05) | 0.02(0.01) | 0.04(0.01) | 0.2(0.02) |
| 50 | 10 | 0.8 | 0.98(0.02) | 0.97(0.03) | 0.01(0.01) | 0.03(0) | 0.15(0.02) |
| 100 | 10 | 0.88 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.11(0.01) |
| 200 | 10 | 0.93 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.07(0.01) |
| 30 | 30 | 0.7 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.16(0.02) |
| 50 | 30 | 0.82 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.13(0.01) |
| 100 | 30 | 0.92 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.09(0.01) |
| 200 | 30 | 0.92 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.06(0.01) |

Table D.4

*Summary of Fit Statistics Across Conditions: Model C, Estimator MLR, ICC$_O$ 0.3 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.49 | 0.91(0.06) | 0.89(0.08) | 0.04(0.02) | 0.07(0.01) | 0.17(0.03) |
| 50 | 5 | 0.73 | 0.96(0.04) | 0.95(0.04) | 0.02(0.02) | 0.05(0.01) | 0.13(0.02) |
| 100 | 5 | 0.91 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0) | 0.09(0.01) |
| 200 | 5 | 0.93 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 0.65 | 0.96(0.04) | 0.95(0.04) | 0.02(0.01) | 0.04(0.01) | 0.15(0.02) |
| 50 | 10 | 0.82 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.03(0) | 0.11(0.02) |
| 100 | 10 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.08(0.01) |
| 200 | 10 | 0.93 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 0.71 | 0.99(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.13(0.02) |
| 50 | 30 | 0.84 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 0.92 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.07(0.01) |
| 200 | 30 | 0.95 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.5

*Summary of Fit Statistics Across Conditions: Model C, Estimator MLR, ICC$_O$ 0.5 and ICC$_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.59 | 0.87(0.1) | 0.84(0.12) | 0.04(0.02) | 0.06(0.01) | 0.19(0.02) |
| 50 | 5 | 0.83 | 0.95(0.05) | 0.94(0.05) | 0.02(0.01) | 0.05(0.01) | 0.15(0.02) |
| 100 | 5 | 0.92 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.03(0) | 0.1(0.01) |
| 200 | 5 | 0.93 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.07(0.01) |
| 30 | 10 | 0.71 | 0.95(0.05) | 0.94(0.06) | 0.02(0.01) | 0.04(0.01) | 0.17(0.02) |
| 50 | 10 | 0.85 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.03(0) | 0.13(0.01) |
| 100 | 10 | 0.93 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.09(0.01) |
| 200 | 10 | 0.93 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 0.7 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.16(0.02) |
| 50 | 30 | 0.89 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.12(0.01) |
| 100 | 30 | 0.93 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.08(0.01) |
| 200 | 30 | 0.94 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.06(0.01) |

Table D.6

*Summary of Fit Statistics Across Conditions: Model C, Estimator MLR, $\text{ICC}_O$ 0.5 and $\text{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.6 | 0.89(0.08) | 0.87(0.1) | 0.04(0.02) | 0.06(0.01) | 0.17(0.03) |
| 50 | 5 | 0.82 | 0.96(0.04) | 0.95(0.05) | 0.02(0.02) | 0.05(0.01) | 0.13(0.02) |
| 100 | 5 | 0.88 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.04(0) | 0.09(0.01) |
| 200 | 5 | 0.94 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.06(0.01) |
| 30 | 10 | 0.71 | 0.95(0.04) | 0.95(0.05) | 0.02(0.01) | 0.04(0.01) | 0.15(0.02) |
| 50 | 10 | 0.83 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.03(0) | 0.12(0.02) |
| 100 | 10 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.08(0.01) |
| 200 | 10 | 0.92 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 0.73 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.14(0.02) |
| 50 | 30 | 0.87 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.11(0.01) |
| 100 | 30 | 0.91 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.08(0.01) |
| 200 | 30 | 0.93 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.7

*Summary of Fit Statistics Across Conditions: Model C, Estimator ULSMV, $\text{ICC}_O$ 0.1 and $\text{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.99(0.02) | 0.99(0.02) | 0(0.01) | 0.07(0.01) | 0.35(0.06) |
| 50 | 5 | 0.99 | 0.99(0.02) | 0.99(0.03) | 0.01(0.01) | 0.06(0.01) | 0.32(0.13) |
| 100 | 5 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.22(0.05) |
| 200 | 5 | 0.95 | 1(0.01) | 1(0.01) | 0(0.01) | 0.03(0) | 0.15(0.03) |
| 30 | 10 | 0.99 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.25(0.06) |
| 50 | 10 | 0.99 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.19(0.04) |
| 100 | 10 | 0.97 | 1(0) | 1(0.01) | 0.01(0.01) | 0.03(0) | 0.13(0.02) |
| 200 | 10 | 0.95 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.09(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0.01) | 0.01(0.01) | 0.03(0) | 0.15(0.02) |
| 50 | 30 | 0.99 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.11(0.01) |
| 100 | 30 | 0.97 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.08(0.01) |
| 200 | 30 | 0.97 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.8

*Summary of Fit Statistics Across Conditions: Model C, Estimator ULSMV, $ICC_O$ 0.1 and $ICC_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.99(0.02) | 0.98(0.03) | 0.01(0.01) | 0.09(0.01) | 0.13(0.03) |
| 50 | 5 | 1 | 0.99(0.02) | 0.98(0.03) | 0.01(0.01) | 0.07(0.01) | 0.1(0.02) |
| 100 | 5 | 0.99 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.07(0.02) |
| 200 | 5 | 0.96 | 0.99(0.01) | 0.99(0.01) | 0(0.01) | 0.03(0) | 0.05(0.01) |
| 30 | 10 | 1 | 0.99(0.02) | 0.99(0.02) | 0(0.01) | 0.06(0.01) | 0.09(0.02) |
| 50 | 10 | 1 | 0.99(0.02) | 0.99(0.02) | 0(0.01) | 0.04(0.01) | 0.07(0.02) |
| 100 | 10 | 0.98 | 0.99(0.01) | 0.99(0.01) | 0(0.01) | 0.03(0) | 0.05(0.01) |
| 200 | 10 | 0.96 | 1(0.01) | 1(0.01) | 0(0) | 0.02(0) | 0.03(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.07(0.02) |
| 50 | 30 | 1 | 1(0) | 1(0.01) | 0(0) | 0.03(0) | 0.05(0.01) |
| 100 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.02(0) | 0.04(0.01) |
| 200 | 30 | 0.98 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.02(0) |

Table D.9

*Summary of Fit Statistics Across Conditions: Model C, Estimator ULSMV, $ICC_O$ 0.3 and $ICC_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.96(0.05) | 0.96(0.06) | 0.01(0.01) | 0.08(0.01) | 0.21(0.04) |
| 50 | 5 | 0.97 | 0.97(0.03) | 0.97(0.04) | 0.01(0.01) | 0.06(0.01) | 0.16(0.03) |
| 100 | 5 | 0.98 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.11(0.01) |
| 200 | 5 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.07(0.01) |
| 30 | 10 | 0.99 | 0.98(0.03) | 0.97(0.04) | 0.01(0.01) | 0.05(0.01) | 0.16(0.02) |
| 50 | 10 | 0.98 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.04(0) | 0.12(0.01) |
| 100 | 10 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 200 | 10 | 0.95 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.03(0) | 0.13(0.01) |
| 50 | 30 | 0.99 | 1(0.01) | 1(0.01) | 0(0) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 0.97 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 0.94 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.10

*Summary of Fit Statistics Across Conditions: Model C, Estimator ULSMV, $\mathrm{ICC}_O$ 0.3 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.96(0.06) | 0.95(0.07) | 0.01(0.01) | 0.09(0.01) | 0.13(0.03) |
| 50 | 5 | 0.97 | 0.97(0.04) | 0.97(0.05) | 0.01(0.01) | 0.07(0.01) | 0.1(0.02) |
| 100 | 5 | 0.97 | 0.99(0.02) | 0.98(0.02) | 0.01(0.01) | 0.05(0.01) | 0.07(0.01) |
| 200 | 5 | 0.94 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.05(0.01) |
| 30 | 10 | 0.99 | 0.98(0.04) | 0.98(0.05) | 0(0.01) | 0.06(0.01) | 0.1(0.02) |
| 50 | 10 | 0.97 | 0.98(0.03) | 0.98(0.03) | 0.01(0.01) | 0.05(0.01) | 0.08(0.01) |
| 100 | 10 | 0.97 | 0.99(0.01) | 0.99(0.02) | 0(0.01) | 0.04(0) | 0.05(0.01) |
| 200 | 10 | 0.95 | 1(0.01) | 0.99(0.01) | 0(0) | 0.02(0) | 0.04(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.04(0.01) | 0.09(0.02) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0.01) | 0.07(0.01) |
| 100 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.02(0) | 0.05(0.01) |
| 200 | 30 | 0.99 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.03(0) |

Table D.11

*Summary of Fit Statistics Across Conditions: Model C, Estimator ULSMV, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.97 | 0.91(0.13) | 0.89(0.15) | 0.01(0.01) | 0.09(0.02) | 0.15(0.02) |
| 50 | 5 | 0.96 | 0.95(0.07) | 0.94(0.08) | 0.01(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 5 | 0.98 | 0.97(0.03) | 0.97(0.04) | 0.01(0.01) | 0.04(0.01) | 0.08(0.01) |
| 200 | 5 | 0.95 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 1 | 0.97(0.06) | 0.97(0.07) | 0(0.01) | 0.06(0.01) | 0.13(0.02) |
| 50 | 10 | 0.99 | 0.97(0.04) | 0.96(0.05) | 0.01(0.01) | 0.04(0.01) | 0.1(0.01) |
| 100 | 10 | 0.95 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.03(0) | 0.07(0.01) |
| 200 | 10 | 0.93 | 0.99(0.01) | 0.99(0.01) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0.01) | 0.12(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.09(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 0.99 | 1(0.01) | 1(0.01) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.12

*Summary of Fit Statistics Across Conditions: Model C, Estimator ULSMV, $ICC_O$ 0.5 and $ICC_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.94(0.1) | 0.92(0.11) | 0.01(0.01) | 0.11(0.02) | 0.12(0.02) |
| 50 | 5 | 0.99 | 0.96(0.06) | 0.95(0.07) | 0.01(0.01) | 0.08(0.01) | 0.09(0.01) |
| 100 | 5 | 0.97 | 0.98(0.03) | 0.97(0.04) | 0.01(0.01) | 0.06(0.01) | 0.06(0.01) |
| 200 | 5 | 0.95 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.04(0.01) |
| 30 | 10 | 1 | 0.99(0.04) | 0.99(0.05) | 0(0) | 0.08(0.01) | 0.1(0.02) |
| 50 | 10 | 0.99 | 0.98(0.04) | 0.98(0.05) | 0(0.01) | 0.06(0.01) | 0.08(0.01) |
| 100 | 10 | 0.97 | 0.98(0.03) | 0.98(0.03) | 0(0.01) | 0.04(0.01) | 0.06(0.01) |
| 200 | 10 | 0.94 | 0.99(0.01) | 0.99(0.02) | 0(0) | 0.03(0) | 0.04(0) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.06(0.01) | 0.09(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.05(0.01) | 0.07(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0.01) | 0.05(0.01) |
| 200 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.04(0) |

Table D.13

*Summary of Fit Statistics Across Conditions: Model C, Estimator WLSMV, $ICC_O$ 0.1 and $ICC_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.99(0.02) | 0.99(0.03) | 0.01(0.01) | 0.07(0.01) | 0.34(0.05) |
| 50 | 5 | 1 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.06(0.01) | 0.29(0.05) |
| 100 | 5 | 0.95 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.22(0.05) |
| 200 | 5 | 0.95 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.14(0.03) |
| 30 | 10 | 0.99 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.26(0.18) |
| 50 | 10 | 0.96 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0) | 0.19(0.04) |
| 100 | 10 | 0.97 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.12(0.02) |
| 200 | 10 | 0.95 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.08(0.01) |
| 30 | 30 | 0.99 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.15(0.02) |
| 50 | 30 | 0.99 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.11(0.01) |
| 100 | 30 | 0.96 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.08(0.01) |
| 200 | 30 | 0.97 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.14

*Summary of Fit Statistics Across Conditions: Model C, Estimator WLSMV, $\text{ICC}_O$ 0.1 and $\text{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.99(0.02) | 0.98(0.03) | 0.01(0.01) | 0.07(0.01) | 0.16(0.04) |
| 50 | 5 | 1 | 0.99(0.02) | 0.98(0.02) | 0.01(0.01) | 0.06(0.01) | 0.12(0.02) |
| 100 | 5 | 0.96 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.08(0.02) |
| 200 | 5 | 0.96 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 1 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.12(0.03) |
| 50 | 10 | 0.99 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.08(0.02) |
| 100 | 10 | 0.97 | 1(0.01) | 1(0.01) | 0(0.01) | 0.03(0) | 0.06(0.01) |
| 200 | 10 | 0.95 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.04(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0.01) | 0(0) | 0.03(0) | 0.08(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.06(0.01) |
| 100 | 30 | 0.98 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.04(0.01) |
| 200 | 30 | 0.97 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.03(0.01) |

Table D.15

*Summary of Fit Statistics Across Conditions: Model C, Estimator WLSMV, $\text{ICC}_O$ 0.3 and $\text{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.98 | 0.97(0.04) | 0.96(0.05) | 0.01(0.01) | 0.07(0.01) | 0.21(0.06) |
| 50 | 5 | 0.96 | 0.98(0.03) | 0.97(0.03) | 0.01(0.01) | 0.06(0.01) | 0.15(0.02) |
| 100 | 5 | 0.97 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.11(0.01) |
| 200 | 5 | 0.94 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.07(0.01) |
| 30 | 10 | 0.97 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.05(0.01) | 0.15(0.02) |
| 50 | 10 | 0.98 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.12(0.01) |
| 100 | 10 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 200 | 10 | 0.95 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.13(0.01) |
| 50 | 30 | 0.99 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 0.98 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 0.96 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.16

*Summary of Fit Statistics Across Conditions: Model C, Estimator WLSMV,* $\text{ICC}_O$ *0.3 and* $\text{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.97 | 0.97(0.05) | 0.96(0.05) | 0.02(0.01) | 0.08(0.01) | 0.15(0.02) |
| 50 | 5 | 0.96 | 0.98(0.03) | 0.98(0.03) | 0.01(0.01) | 0.06(0.01) | 0.11(0.02) |
| 100 | 5 | 0.97 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.08(0.01) |
| 200 | 5 | 0.95 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.05(0.01) |
| 30 | 10 | 0.99 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.12(0.02) |
| 50 | 10 | 0.97 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.09(0.01) |
| 100 | 10 | 0.97 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 200 | 10 | 0.96 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.04(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.1(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.08(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.05(0.01) |
| 200 | 30 | 0.99 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.04(0.01) |

Table D.17

*Summary of Fit Statistics Across Conditions: Model C, Estimator WLSMV,* $\text{ICC}_O$ *0.5 and* $\text{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.96 | 0.95(0.05) | 0.94(0.06) | 0.02(0.01) | 0.08(0.01) | 0.16(0.02) |
| 50 | 5 | 0.97 | 0.98(0.03) | 0.97(0.03) | 0.01(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 5 | 0.99 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.08(0.01) |
| 200 | 5 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 0.99 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.13(0.02) |
| 50 | 10 | 0.99 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0.01) | 0.1(0.01) |
| 100 | 10 | 0.97 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.07(0.01) |
| 200 | 10 | 0.96 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.12(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.09(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.18

*Summary of Fit Statistics Across Conditions: Model C, Estimator WLSMV,* $\mathrm{ICC}_O$ *0.5 and* $\mathrm{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.97(0.04) | 0.96(0.05) | 0.01(0.01) | 0.08(0.01) | 0.14(0.02) |
| 50 | 5 | 0.98 | 0.98(0.03) | 0.97(0.03) | 0.01(0.01) | 0.06(0.01) | 0.1(0.01) |
| 100 | 5 | 0.95 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.07(0.01) |
| 200 | 5 | 0.95 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.05(0.01) |
| 30 | 10 | 1 | 0.99(0.01) | 0.99(0.02) | 0(0.01) | 0.05(0.01) | 0.12(0.02) |
| 50 | 10 | 0.99 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0.01) | 0.09(0.01) |
| 100 | 10 | 0.96 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 200 | 10 | 0.95 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.11(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.08(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.06(0.01) |
| 200 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.04(0.01) |

Table D.19

*Summary of Fit Statistics Across Conditions: Model M1, Estimator MLR,* $\mathrm{ICC}_O$ *0.1 and* $\mathrm{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.11 | 0.78(0.14) | 0.74(0.16) | 0.07(0.02) | 0.08(0.03) | 0.32(0.06) |
| 50 | 5 | 0.15 | 0.88(0.07) | 0.85(0.08) | 0.05(0.01) | 0.06(0.01) | 0.29(0.05) |
| 100 | 5 | 0.05 | 0.91(0.03) | 0.89(0.04) | 0.04(0.01) | 0.05(0.01) | 0.24(0.04) |
| 200 | 5 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.19(0.03) |
| 30 | 10 | 0.04 | 0.86(0.08) | 0.83(0.09) | 0.05(0.02) | 0.06(0.01) | 0.25(0.04) |
| 50 | 10 | 0.01 | 0.9(0.03) | 0.88(0.04) | 0.04(0.01) | 0.05(0.01) | 0.21(0.03) |
| 100 | 10 | 0 | 0.92(0.02) | 0.9(0.03) | 0.03(0.01) | 0.05(0) | 0.16(0.02) |
| 200 | 10 | 0 | 0.92(0.01) | 0.91(0.02) | 0.03(0) | 0.04(0) | 0.12(0.02) |
| 30 | 30 | 0 | 0.9(0.03) | 0.88(0.04) | 0.04(0.01) | 0.05(0.01) | 0.18(0.03) |
| 50 | 30 | 0 | 0.91(0.04) | 0.89(0.05) | 0.04(0.01) | 0.04(0) | 0.14(0.02) |
| 100 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.1(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.08(0.01) |

Table D.20

*Summary of Fit Statistics Across Conditions: Model M1, Estimator MLR, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.12 | 0.88(0.06) | 0.85(0.07) | 0.06(0.02) | 0.08(0.01) | 0.16(0.03) |
| 50 | 5 | 0.15 | 0.92(0.04) | 0.91(0.04) | 0.05(0.01) | 0.07(0.01) | 0.13(0.03) |
| 100 | 5 | 0.02 | 0.95(0.02) | 0.93(0.02) | 0.04(0.01) | 0.06(0.01) | 0.1(0.02) |
| 200 | 5 | 0 | 0.95(0.01) | 0.94(0.02) | 0.03(0.01) | 0.05(0.01) | 0.08(0.01) |
| 30 | 10 | 0.05 | 0.9(0.04) | 0.88(0.05) | 0.05(0.01) | 0.06(0.01) | 0.13(0.03) |
| 50 | 10 | 0.01 | 0.92(0.02) | 0.91(0.03) | 0.04(0.01) | 0.06(0.01) | 0.1(0.02) |
| 100 | 10 | 0 | 0.94(0.01) | 0.92(0.02) | 0.04(0) | 0.05(0) | 0.08(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.93(0.01) | 0.04(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 0 | 0.91(0.02) | 0.89(0.02) | 0.04(0.01) | 0.05(0.01) | 0.11(0.02) |
| 50 | 30 | 0 | 0.92(0.01) | 0.91(0.02) | 0.04(0) | 0.05(0) | 0.09(0.02) |
| 100 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.04(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.04(0) | 0.04(0) | 0.06(0.01) |

Table D.21

*Summary of Fit Statistics Across Conditions: Model M1, Estimator MLR, $\mathrm{ICC}_O$ 0.3 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.27 | 0.8(0.12) | 0.76(0.14) | 0.05(0.02) | 0.07(0.02) | 0.23(0.03) |
| 50 | 5 | 0.36 | 0.89(0.06) | 0.87(0.08) | 0.04(0.01) | 0.06(0.01) | 0.19(0.02) |
| 100 | 5 | 0.19 | 0.92(0.04) | 0.9(0.05) | 0.03(0.01) | 0.05(0.01) | 0.14(0.02) |
| 200 | 5 | 0.01 | 0.92(0.02) | 0.9(0.03) | 0.03(0.01) | 0.05(0.01) | 0.1(0.01) |
| 30 | 10 | 0.14 | 0.87(0.06) | 0.84(0.08) | 0.04(0.01) | 0.06(0.01) | 0.19(0.03) |
| 50 | 10 | 0.07 | 0.9(0.04) | 0.88(0.05) | 0.04(0.01) | 0.05(0.01) | 0.15(0.02) |
| 100 | 10 | 0.01 | 0.91(0.03) | 0.89(0.03) | 0.03(0.01) | 0.05(0.01) | 0.1(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.07(0.01) |
| 30 | 30 | 0 | 0.9(0.03) | 0.88(0.03) | 0.04(0.01) | 0.05(0.01) | 0.16(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.89(0.02) | 0.03(0) | 0.04(0) | 0.12(0.01) |
| 100 | 30 | 0 | 0.91(0.01) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.09(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.06(0.01) |

Table D.22

*Summary of Fit Statistics Across Conditions: Model M1, Estimator MLR, $ICC_O$ 0.3 and $ICC_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.26 | 0.85(0.08) | 0.83(0.09) | 0.05(0.02) | 0.08(0.01) | 0.17(0.03) |
| 50 | 5 | 0.31 | 0.91(0.05) | 0.89(0.06) | 0.04(0.01) | 0.07(0.01) | 0.14(0.02) |
| 100 | 5 | 0.1 | 0.93(0.03) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.1(0.01) |
| 200 | 5 | 0.01 | 0.94(0.02) | 0.93(0.02) | 0.03(0) | 0.05(0.01) | 0.07(0.01) |
| 30 | 10 | 0.12 | 0.89(0.04) | 0.87(0.05) | 0.04(0.01) | 0.06(0.01) | 0.15(0.02) |
| 50 | 10 | 0.04 | 0.92(0.03) | 0.9(0.04) | 0.04(0.01) | 0.05(0.01) | 0.12(0.02) |
| 100 | 10 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0.01) | 0.09(0.01) |
| 200 | 10 | 0 | 0.93(0.01) | 0.91(0.01) | 0.03(0) | 0.04(0) | 0.07(0.01) |
| 30 | 30 | 0 | 0.91(0.02) | 0.89(0.03) | 0.04(0.01) | 0.05(0) | 0.13(0.02) |
| 50 | 30 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.11(0.02) |
| 100 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.03(0) | 0.04(0) | 0.07(0.01) |

Table D.23

*Summary of Fit Statistics Across Conditions: Model M1, Estimator MLR, $ICC_O$ 0.5 and $ICC_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.34 | 0.8(0.1) | 0.76(0.13) | 0.05(0.02) | 0.07(0.01) | 0.19(0.02) |
| 50 | 5 | 0.42 | 0.89(0.07) | 0.87(0.08) | 0.03(0.01) | 0.06(0.01) | 0.15(0.02) |
| 100 | 5 | 0.23 | 0.91(0.04) | 0.89(0.05) | 0.03(0.01) | 0.05(0.01) | 0.1(0.01) |
| 200 | 5 | 0.01 | 0.91(0.03) | 0.9(0.03) | 0.03(0.01) | 0.05(0.01) | 0.07(0.01) |
| 30 | 10 | 0.22 | 0.88(0.06) | 0.85(0.08) | 0.04(0.01) | 0.06(0.01) | 0.17(0.02) |
| 50 | 10 | 0.09 | 0.9(0.04) | 0.88(0.05) | 0.03(0.01) | 0.05(0.01) | 0.13(0.01) |
| 100 | 10 | 0 | 0.91(0.03) | 0.9(0.03) | 0.03(0.01) | 0.04(0) | 0.09(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.06(0.01) |
| 30 | 30 | 0 | 0.9(0.03) | 0.88(0.03) | 0.04(0) | 0.05(0.01) | 0.15(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.12(0.01) |
| 100 | 30 | 0 | 0.91(0.03) | 0.9(0.03) | 0.03(0) | 0.04(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.06(0.01) |

Table D.24

*Summary of Fit Statistics Across Conditions: Model M1, Estimator MLR,* $\mathrm{ICC}_O$ *0.5 and* $\mathrm{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.37 | 0.84(0.09) | 0.8(0.11) | 0.05(0.02) | 0.07(0.01) | 0.17(0.02) |
| 50 | 5 | 0.4 | 0.9(0.06) | 0.88(0.07) | 0.04(0.01) | 0.06(0.01) | 0.13(0.02) |
| 100 | 5 | 0.15 | 0.92(0.03) | 0.9(0.04) | 0.03(0.01) | 0.05(0.01) | 0.09(0.01) |
| 200 | 5 | 0 | 0.93(0.02) | 0.91(0.03) | 0.03(0) | 0.05(0.01) | 0.07(0.01) |
| 30 | 10 | 0.2 | 0.88(0.07) | 0.86(0.08) | 0.04(0.01) | 0.06(0.01) | 0.15(0.02) |
| 50 | 10 | 0.09 | 0.91(0.04) | 0.89(0.05) | 0.03(0.01) | 0.05(0.01) | 0.12(0.02) |
| 100 | 10 | 0 | 0.92(0.02) | 0.9(0.03) | 0.03(0) | 0.04(0) | 0.08(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.91(0.02) | 0.03(0) | 0.04(0) | 0.06(0.01) |
| 30 | 30 | 0 | 0.91(0.03) | 0.89(0.03) | 0.03(0.01) | 0.04(0) | 0.14(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.11(0.01) |
| 100 | 30 | 0 | 0.92(0.01) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.06(0.01) |

Table D.25

*Summary of Fit Statistics Across Conditions: Model M1, Estimator ULSMV,* $\mathrm{ICC}_O$ *0.1 and* $\mathrm{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.96(0.05) | 0.95(0.06) | 0.02(0.02) | 0.09(0.01) | 0.52(0.68) |
| 50 | 5 | 0.65 | 0.93(0.05) | 0.91(0.06) | 0.03(0.01) | 0.07(0.01) | 0.32(0.14) |
| 100 | 5 | 0.09 | 0.92(0.03) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.22(0.05) |
| 200 | 5 | 0 | 0.92(0.02) | 0.91(0.03) | 0.04(0.01) | 0.05(0.01) | 0.14(0.03) |
| 30 | 10 | 0.44 | 0.93(0.04) | 0.92(0.04) | 0.03(0.01) | 0.07(0.01) | 0.26(0.12) |
| 50 | 10 | 0.05 | 0.92(0.03) | 0.91(0.04) | 0.04(0.01) | 0.06(0.01) | 0.19(0.04) |
| 100 | 10 | 0 | 0.92(0.02) | 0.91(0.03) | 0.04(0.01) | 0.05(0.01) | 0.13(0.02) |
| 200 | 10 | 0 | 0.92(0.01) | 0.91(0.02) | 0.04(0) | 0.05(0) | 0.09(0.01) |
| 30 | 30 | 0.01 | 0.93(0.02) | 0.92(0.02) | 0.03(0.01) | 0.05(0.01) | 0.15(0.02) |
| 50 | 30 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.11(0.01) |
| 100 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.04(0) | 0.05(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.04(0) | 0.05(0) | 0.06(0.01) |

Table D.26

*Summary of Fit Statistics Across Conditions: Model M1, Estimator ULSMV,* $\mathrm{ICC}_O$ *0.1 and* $\mathrm{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.97(0.05) | 0.96(0.06) | 0.01(0.01) | 0.09(0.02) | 0.13(0.04) |
| 50 | 5 | 0.55 | 0.91(0.06) | 0.89(0.07) | 0.03(0.01) | 0.09(0.01) | 0.1(0.03) |
| 100 | 5 | 0.31 | 0.92(0.04) | 0.9(0.05) | 0.03(0.01) | 0.07(0.01) | 0.07(0.01) |
| 200 | 5 | 0 | 0.95(NA) | 0.94(NA) | 0.02(NA) | 0.06(NA) | 0.04(NA) |
| 30 | 10 | 0.78 | 0.91(0.08) | 0.89(0.1) | 0.02(0.01) | 0.08(0.01) | 0.09(0.02) |
| 50 | 10 | 0.2 | 0.88(0.05) | 0.85(0.06) | 0.03(0.01) | 0.07(0.01) | 0.07(0.02) |
| 100 | 10 | 0 | 0.9(NA) | 0.88(NA) | 0.03(NA) | 0.05(NA) | 0.06(NA) |
| 30 | 30 | 0.5 | 0.92(0.05) | 0.91(0.05) | 0.02(0.01) | 0.06(0) | 0.08(0.03) |
| 50 | 30 | 0 | 0.82(NA) | 0.78(NA) | 0.02(NA) | 0.06(NA) | 0.06(NA) |

Table D.27

*Summary of Fit Statistics Across Conditions: Model M1, Estimator ULSMV,* $\mathrm{ICC}_O$ *0.3 and* $\mathrm{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.93 | 0.91(0.08) | 0.9(0.09) | 0.02(0.01) | 0.09(0.01) | 0.21(0.04) |
| 50 | 5 | 0.76 | 0.93(0.05) | 0.91(0.06) | 0.02(0.01) | 0.07(0.01) | 0.15(0.03) |
| 100 | 5 | 0.3 | 0.93(0.04) | 0.92(0.04) | 0.03(0.01) | 0.06(0.01) | 0.1(0.01) |
| 200 | 5 | 0.02 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 30 | 10 | 0.82 | 0.92(0.05) | 0.91(0.06) | 0.02(0.01) | 0.07(0.01) | 0.15(0.02) |
| 50 | 10 | 0.38 | 0.93(0.03) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 10 | 0.02 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.08(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.93(0.02) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 0.93 | 0.97(0.03) | 0.96(0.03) | 0.01(0.01) | 0.06(0.01) | 0.13(0.01) |
| 50 | 30 | 0.3 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.05(0) | 0.1(0.01) |
| 100 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.02(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.02(0) | 0.05(0) | 0.05(0.01) |

Table D.28

*Summary of Fit Statistics Across Conditions: Model M1, Estimator ULSMV, $\mathrm{ICC}_O$ 0.3 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.97 | 0.91(0.09) | 0.9(0.11) | 0.02(0.01) | 0.1(0.01) | 0.13(0.03) |
| 50 | 5 | 0.79 | 0.91(0.07) | 0.89(0.08) | 0.02(0.01) | 0.09(0.01) | 0.1(0.02) |
| 100 | 5 | 0.4 | 0.92(0.04) | 0.9(0.05) | 0.03(0.01) | 0.07(0.01) | 0.07(0.01) |
| 200 | 5 | 0.01 | 0.92(0.03) | 0.9(0.03) | 0.03(0) | 0.06(0.01) | 0.05(0.01) |
| 30 | 10 | 0.95 | 0.93(0.07) | 0.91(0.09) | 0.02(0.01) | 0.08(0.01) | 0.11(0.02) |
| 50 | 10 | 0.67 | 0.91(0.06) | 0.89(0.07) | 0.02(0.01) | 0.07(0.01) | 0.08(0.01) |
| 100 | 10 | 0.08 | 0.91(0.04) | 0.89(0.04) | 0.02(0) | 0.06(0.01) | 0.06(0.01) |
| 200 | 10 | 0 | 0.9(0.02) | 0.88(0.03) | 0.02(0) | 0.06(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.07(0.01) | 0.09(0.01) |
| 50 | 30 | 0.98 | 0.98(0.03) | 0.98(0.03) | 0(0) | 0.06(0.01) | 0.07(0.01) |
| 100 | 30 | 0.11 | 0.93(0.03) | 0.92(0.03) | 0.01(0) | 0.06(0) | 0.05(0.01) |
| 200 | 30 | 0 | 0.9(0.02) | 0.88(0.02) | 0.02(0) | 0.05(0) | 0.04(0) |

Table D.29

*Summary of Fit Statistics Across Conditions: Model M1, Estimator ULSMV, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.97 | 0.88(0.13) | 0.85(0.16) | 0.02(0.01) | 0.1(0.02) | 0.15(0.02) |
| 50 | 5 | 0.94 | 0.91(0.08) | 0.89(0.1) | 0.02(0.01) | 0.08(0.01) | 0.11(0.01) |
| 100 | 5 | 0.77 | 0.93(0.05) | 0.92(0.06) | 0.02(0.01) | 0.06(0.01) | 0.08(0.01) |
| 200 | 5 | 0.34 | 0.94(0.03) | 0.93(0.03) | 0.02(0) | 0.06(0.01) | 0.06(0.01) |
| 30 | 10 | 0.99 | 0.95(0.07) | 0.94(0.09) | 0.01(0.01) | 0.07(0.01) | 0.13(0.01) |
| 50 | 10 | 0.93 | 0.94(0.06) | 0.92(0.07) | 0.01(0.01) | 0.06(0.01) | 0.1(0.01) |
| 100 | 10 | 0.66 | 0.94(0.04) | 0.93(0.04) | 0.01(0.01) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0.12 | 0.94(0.02) | 0.93(0.02) | 0.02(0) | 0.05(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.06(0.01) | 0.12(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0.01) | 0(0) | 0.05(0) | 0.09(0.01) |
| 100 | 30 | 0.99 | 0.98(0.02) | 0.98(0.02) | 0(0) | 0.05(0) | 0.06(0.01) |
| 200 | 30 | 0.23 | 0.96(0.01) | 0.95(0.02) | 0.01(0) | 0.05(0) | 0.05(0) |

Table D.30

*Summary of Fit Statistics Across Conditions: Model M1, Estimator ULSMV, $ICC_O$ 0.5 and $ICC_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.91(0.11) | 0.89(0.13) | 0.01(0.01) | 0.12(0.02) | 0.12(0.02) |
| 50 | 5 | 0.95 | 0.92(0.08) | 0.9(0.1) | 0.02(0.01) | 0.09(0.01) | 0.09(0.01) |
| 100 | 5 | 0.73 | 0.93(0.05) | 0.91(0.06) | 0.02(0.01) | 0.08(0.01) | 0.06(0.01) |
| 200 | 5 | 0.27 | 0.93(0.03) | 0.92(0.03) | 0.02(0) | 0.07(0.01) | 0.05(0.01) |
| 30 | 10 | 1 | 0.98(0.06) | 0.97(0.07) | 0(0.01) | 0.09(0.01) | 0.1(0.02) |
| 50 | 10 | 0.97 | 0.95(0.06) | 0.94(0.07) | 0.01(0.01) | 0.08(0.01) | 0.08(0.01) |
| 100 | 10 | 0.68 | 0.94(0.04) | 0.92(0.05) | 0.01(0.01) | 0.07(0.01) | 0.06(0.01) |
| 200 | 10 | 0.12 | 0.93(0.03) | 0.92(0.03) | 0.02(0) | 0.06(0.01) | 0.04(0) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.08(0.01) | 0.09(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.07(0.01) | 0.07(0.01) |
| 100 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.06(0.01) | 0.05(0.01) |
| 200 | 30 | 0.68 | 0.97(0.02) | 0.96(0.02) | 0.01(0) | 0.06(0) | 0.04(0) |

Table D.31

*Summary of Fit Statistics Across Conditions: Model M1, Estimator WLSMV, $ICC_O$ 0.1 and $ICC_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.9 | 0.94(0.06) | 0.93(0.07) | 0.02(0.02) | 0.09(0.01) | 0.33(0.04) |
| 50 | 5 | 0.68 | 0.93(0.05) | 0.91(0.06) | 0.03(0.01) | 0.07(0.01) | 0.28(0.05) |
| 100 | 5 | 0.2 | 0.93(0.04) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.21(0.04) |
| 200 | 5 | 0 | 0.93(0.02) | 0.91(0.03) | 0.03(0.01) | 0.05(0.01) | 0.14(0.02) |
| 30 | 10 | 0.51 | 0.93(0.04) | 0.92(0.05) | 0.03(0.01) | 0.07(0.01) | 0.3(0.69) |
| 50 | 10 | 0.15 | 0.93(0.03) | 0.91(0.03) | 0.03(0.01) | 0.06(0.01) | 0.19(0.04) |
| 100 | 10 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0.01) | 0.12(0.02) |
| 200 | 10 | 0 | 0.93(0.01) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.09(0.01) |
| 30 | 30 | 0.01 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.05(0.01) | 0.15(0.02) |
| 50 | 30 | 0 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.05(0) | 0.11(0.01) |
| 100 | 30 | 0 | 0.93(0.01) | 0.91(0.01) | 0.03(0) | 0.05(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.93(0.01) | 0.91(0.01) | 0.04(0) | 0.05(0) | 0.06(0.01) |

Table D.32

*Summary of Fit Statistics Across Conditions: Model M1, Estimator WLSMV,* $\mathrm{ICC}_O$ *0.1 and* $\mathrm{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.95(0.05) | 0.94(0.06) | 0.02(0.02) | 0.08(0.01) | 0.14(0.05) |
| 50 | 5 | 0 | 0.86(0.01) | 0.83(0.01) | 0.04(0) | 0.09(0.01) | 0.14(0.07) |
| 100 | 5 | 0.4 | 0.93(0.04) | 0.91(0.05) | 0.03(0.01) | 0.06(0.01) | 0.08(0.02) |
| 30 | 10 | 0.4 | 0.91(0.06) | 0.9(0.07) | 0.03(0.01) | 0.07(0.01) | 0.14(0.03) |
| 50 | 10 | 0 | 0.91(0.04) | 0.89(0.05) | 0.03(0.01) | 0.07(0.02) | 0.1(0.03) |
| 30 | 30 | 0.5 | 0.96(0.03) | 0.95(0.04) | 0.02(0.01) | 0.05(0.01) | 0.08(0.01) |

Table D.33

*Summary of Fit Statistics Across Conditions: Model M1, Estimator WLSMV,* $\mathrm{ICC}_O$ *0.3 and* $\mathrm{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.87 | 0.92(0.07) | 0.9(0.08) | 0.03(0.02) | 0.09(0.01) | 0.2(0.03) |
| 50 | 5 | 0.59 | 0.92(0.05) | 0.91(0.06) | 0.03(0.01) | 0.07(0.01) | 0.15(0.02) |
| 100 | 5 | 0.21 | 0.93(0.03) | 0.92(0.04) | 0.03(0.01) | 0.06(0.01) | 0.1(0.01) |
| 200 | 5 | 0 | 0.93(0.02) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.07(0.01) |
| 30 | 10 | 0.51 | 0.92(0.05) | 0.91(0.05) | 0.03(0.01) | 0.07(0.01) | 0.15(0.02) |
| 50 | 10 | 0.13 | 0.93(0.03) | 0.91(0.03) | 0.03(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 10 | 0.01 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.06(0.01) | 0.08(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.92(0.02) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 0.12 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.06(0.01) | 0.13(0.01) |
| 50 | 30 | 0 | 0.94(0.01) | 0.93(0.02) | 0.03(0) | 0.05(0) | 0.1(0.01) |
| 100 | 30 | 0 | 0.94(0.01) | 0.92(0.01) | 0.03(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.92(0.01) | 0.03(0) | 0.05(0) | 0.05(0.01) |

Table D.34

*Summary of Fit Statistics Across Conditions: Model M1, Estimator WLSMV, ICC$_O$ 0.3 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.81 | 0.91(0.07) | 0.9(0.08) | 0.03(0.02) | 0.09(0.01) | 0.15(0.03) |
| 50 | 5 | 0.56 | 0.92(0.06) | 0.9(0.07) | 0.03(0.01) | 0.08(0.01) | 0.11(0.02) |
| 100 | 5 | 0.18 | 0.93(0.04) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.08(0.01) |
| 200 | 5 | 0 | 0.93(0.02) | 0.91(0.03) | 0.03(0) | 0.06(0.01) | 0.06(0.01) |
| 30 | 10 | 0.56 | 0.93(0.04) | 0.91(0.05) | 0.03(0.01) | 0.07(0.01) | 0.13(0.02) |
| 50 | 10 | 0.15 | 0.92(0.03) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.09(0.01) |
| 100 | 10 | 0 | 0.92(0.03) | 0.9(0.03) | 0.03(0.01) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.05(0) | 0.05(0.01) |
| 30 | 30 | 0.8 | 0.97(0.02) | 0.97(0.02) | 0.01(0.01) | 0.06(0.01) | 0.1(0.01) |
| 50 | 30 | 0.04 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.05(0.01) | 0.08(0.01) |
| 100 | 30 | 0 | 0.93(0.01) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.05(0) | 0.05(0.01) |

Table D.35

*Summary of Fit Statistics Across Conditions: Model M1, Estimator WLSMV, ICC$_O$ 0.3 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.82 | 0.91(0.07) | 0.89(0.08) | 0.03(0.02) | 0.09(0.01) | 0.15(0.02) |
| 50 | 5 | 0.71 | 0.93(0.05) | 0.91(0.06) | 0.03(0.01) | 0.08(0.01) | 0.11(0.01) |
| 100 | 5 | 0.28 | 0.94(0.03) | 0.92(0.04) | 0.03(0.01) | 0.06(0.01) | 0.08(0.01) |
| 200 | 5 | 0.01 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.06(0.01) |
| 30 | 10 | 0.8 | 0.95(0.04) | 0.93(0.05) | 0.02(0.01) | 0.07(0.01) | 0.13(0.02) |
| 50 | 10 | 0.3 | 0.94(0.03) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.1(0.01) |
| 100 | 10 | 0 | 0.94(0.02) | 0.92(0.02) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.92(0.02) | 0.03(0) | 0.05(0) | 0.05(0.01) |
| 30 | 30 | 0.99 | 0.99(0.01) | 0.99(0.02) | 0(0.01) | 0.06(0.01) | 0.12(0.01) |
| 50 | 30 | 0.37 | 0.97(0.01) | 0.96(0.02) | 0.01(0) | 0.05(0.01) | 0.09(0.01) |
| 100 | 30 | 0 | 0.95(0.01) | 0.94(0.01) | 0.02(0) | 0.05(0) | 0.06(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.03(0) | 0.05(0) | 0.05(0) |

Table D.36

*Summary of Fit Statistics Across Conditions: Model M1, Estimator WLSMV, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.89 | 0.93(0.06) | 0.91(0.07) | 0.03(0.01) | 0.09(0.01) | 0.14(0.02) |
| 50 | 5 | 0.66 | 0.93(0.05) | 0.91(0.06) | 0.03(0.01) | 0.08(0.01) | 0.1(0.01) |
| 100 | 5 | 0.2 | 0.93(0.03) | 0.91(0.04) | 0.03(0.01) | 0.07(0.01) | 0.07(0.01) |
| 200 | 5 | 0 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.05(0.01) |
| 30 | 10 | 0.88 | 0.96(0.04) | 0.95(0.05) | 0.02(0.01) | 0.07(0.01) | 0.12(0.02) |
| 50 | 10 | 0.4 | 0.94(0.03) | 0.93(0.04) | 0.03(0.01) | 0.06(0.01) | 0.09(0.01) |
| 100 | 10 | 0.01 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0.01) | 0(0) | 0.06(0.01) | 0.11(0.01) |
| 50 | 30 | 0.86 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.06(0.01) | 0.09(0.01) |
| 100 | 30 | 0 | 0.95(0.01) | 0.94(0.01) | 0.02(0) | 0.05(0) | 0.06(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.03(0) | 0.05(0) | 0.05(0.01) |

Table D.37

*Summary of Fit Statistics Across Conditions: Model M2, Estimator MLR, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.35 | 0.87(0.11) | 0.84(0.12) | 0.05(0.02) | 0.06(0.03) | 0.33(0.06) |
| 50 | 5 | 0.52 | 0.94(0.04) | 0.93(0.05) | 0.03(0.02) | 0.05(0.01) | 0.28(0.04) |
| 100 | 5 | 0.69 | 0.98(0.02) | 0.97(0.03) | 0.02(0.01) | 0.03(0) | 0.23(0.03) |
| 200 | 5 | 0.84 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.18(0.03) |
| 30 | 10 | 0.42 | 0.94(0.04) | 0.92(0.05) | 0.03(0.01) | 0.04(0.01) | 0.26(0.04) |
| 50 | 10 | 0.6 | 0.97(0.02) | 0.96(0.03) | 0.02(0.01) | 0.03(0) | 0.22(0.03) |
| 100 | 10 | 0.81 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.17(0.02) |
| 200 | 10 | 0.79 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.12(0.02) |
| 30 | 30 | 0.56 | 0.98(0.01) | 0.98(0.02) | 0.02(0.01) | 0.02(0) | 0.2(0.03) |
| 50 | 30 | 0.72 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.15(0.02) |
| 100 | 30 | 0.74 | 1(0) | 0.99(0) | 0.01(0) | 0.01(0) | 0.12(0.01) |
| 200 | 30 | 0.59 | 1(0) | 1(0) | 0.01(0) | 0.01(0) | 0.09(0.01) |

Table D.38

*Summary of Fit Statistics Across Conditions: Model M2, Estimator MLR, ICC$_O$ 0.1 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.2 | 0.89(0.06) | 0.87(0.07) | 0.06(0.02) | 0.07(0.01) | 0.21(0.05) |
| 50 | 5 | 0.2 | 0.93(0.03) | 0.91(0.04) | 0.04(0.01) | 0.05(0.01) | 0.18(0.03) |
| 100 | 5 | 0.09 | 0.95(0.02) | 0.94(0.03) | 0.03(0.01) | 0.04(0) | 0.16(0.03) |
| 200 | 5 | 0 | 0.96(0.01) | 0.95(0.02) | 0.03(0.01) | 0.03(0) | 0.14(0.02) |
| 30 | 10 | 0.15 | 0.92(0.04) | 0.91(0.04) | 0.04(0.01) | 0.04(0.01) | 0.17(0.04) |
| 50 | 10 | 0.09 | 0.94(0.02) | 0.93(0.03) | 0.03(0.01) | 0.04(0) | 0.16(0.03) |
| 100 | 10 | 0 | 0.96(0.01) | 0.95(0.01) | 0.03(0) | 0.03(0) | 0.14(0.02) |
| 200 | 10 | 0 | 0.96(0.01) | 0.95(0.01) | 0.03(0) | 0.02(0) | 0.13(0.01) |
| 30 | 30 | 0.08 | 0.96(0.02) | 0.95(0.02) | 0.03(0.01) | 0.02(0) | 0.16(0.03) |
| 50 | 30 | 0.03 | 0.97(0.01) | 0.97(0.01) | 0.02(0) | 0.02(0) | 0.15(0.02) |
| 100 | 30 | 0 | 0.98(0.01) | 0.97(0.01) | 0.02(0) | 0.01(0) | 0.14(0.02) |
| 200 | 30 | 0 | 0.98(0) | 0.97(0) | 0.02(0) | 0.01(0) | 0.13(0.01) |

Table D.39

*Summary of Fit Statistics Across Conditions: Model M2, Estimator MLR, ICC$_O$ 0.3 and ICC$_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.47 | 0.87(0.09) | 0.85(0.11) | 0.04(0.02) | 0.06(0.01) | 0.23(0.03) |
| 50 | 5 | 0.78 | 0.95(0.05) | 0.94(0.06) | 0.02(0.02) | 0.05(0.01) | 0.19(0.02) |
| 100 | 5 | 0.88 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.03(0) | 0.14(0.02) |
| 200 | 5 | 0.87 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.1(0.01) |
| 30 | 10 | 0.63 | 0.94(0.04) | 0.93(0.05) | 0.02(0.01) | 0.04(0.01) | 0.2(0.02) |
| 50 | 10 | 0.77 | 0.98(0.02) | 0.97(0.03) | 0.01(0.01) | 0.03(0) | 0.16(0.02) |
| 100 | 10 | 0.87 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.11(0.01) |
| 200 | 10 | 0.88 | 1(0) | 1(0.01) | 0.01(0) | 0.02(0) | 0.08(0.01) |
| 30 | 30 | 0.71 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.16(0.02) |
| 50 | 30 | 0.8 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.13(0.01) |
| 100 | 30 | 0.88 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.09(0.01) |
| 200 | 30 | 0.88 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.07(0.01) |

Table D.40

*Summary of Fit Statistics Across Conditions: Model M2, Estimator MLR, $ICC_O$ 0.3 and $ICC_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.37 | 0.88(0.07) | 0.86(0.08) | 0.05(0.02) | 0.07(0.01) | 0.19(0.03) |
| 50 | 5 | 0.52 | 0.94(0.04) | 0.93(0.05) | 0.03(0.02) | 0.05(0.01) | 0.16(0.02) |
| 100 | 5 | 0.46 | 0.96(0.02) | 0.96(0.03) | 0.02(0.01) | 0.04(0) | 0.12(0.02) |
| 200 | 5 | 0.14 | 0.97(0.01) | 0.96(0.02) | 0.02(0.01) | 0.03(0) | 0.11(0.01) |
| 30 | 10 | 0.47 | 0.94(0.04) | 0.92(0.05) | 0.03(0.01) | 0.04(0.01) | 0.17(0.03) |
| 50 | 10 | 0.5 | 0.96(0.02) | 0.96(0.03) | 0.02(0.01) | 0.03(0) | 0.14(0.02) |
| 100 | 10 | 0.26 | 0.97(0.01) | 0.97(0.02) | 0.02(0.01) | 0.02(0) | 0.11(0.01) |
| 200 | 10 | 0.02 | 0.98(0.01) | 0.97(0.01) | 0.02(0) | 0.02(0) | 0.1(0.01) |
| 30 | 30 | 0.49 | 0.98(0.02) | 0.97(0.02) | 0.02(0.01) | 0.02(0) | 0.16(0.02) |
| 50 | 30 | 0.43 | 0.98(0.01) | 0.98(0.01) | 0.01(0.01) | 0.02(0) | 0.13(0.02) |
| 100 | 30 | 0.19 | 0.99(0.01) | 0.99(0.01) | 0.01(0) | 0.01(0) | 0.11(0.01) |
| 200 | 30 | 0 | 0.99(0) | 0.99(0) | 0.01(0) | 0.01(0) | 0.1(0.01) |

Table D.41

*Summary of Fit Statistics Across Conditions: Model M2, Estimator MLR, $ICC_O$ 0.5 and $ICC_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.57 | 0.87(0.1) | 0.84(0.12) | 0.04(0.02) | 0.06(0.01) | 0.19(0.02) |
| 50 | 5 | 0.83 | 0.95(0.05) | 0.94(0.05) | 0.02(0.01) | 0.05(0.01) | 0.15(0.02) |
| 100 | 5 | 0.92 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.03(0) | 0.11(0.01) |
| 200 | 5 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.07(0.01) |
| 30 | 10 | 0.75 | 0.95(0.04) | 0.94(0.05) | 0.02(0.01) | 0.04(0.01) | 0.17(0.02) |
| 50 | 10 | 0.85 | 0.98(0.02) | 0.97(0.03) | 0.01(0.01) | 0.03(0) | 0.13(0.01) |
| 100 | 10 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.09(0.01) |
| 200 | 10 | 0.92 | 1(0) | 1(0.01) | 0(0) | 0.02(0) | 0.07(0.01) |
| 30 | 30 | 0.69 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.16(0.02) |
| 50 | 30 | 0.87 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.12(0.01) |
| 100 | 30 | 0.91 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.08(0.01) |
| 200 | 30 | 0.94 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.06(0.01) |

Table D.42

*Summary of Fit Statistics Across Conditions: Model M2, Estimator MLR, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.51 | 0.88(0.09) | 0.85(0.1) | 0.04(0.02) | 0.06(0.01) | 0.18(0.03) |
| 50 | 5 | 0.73 | 0.95(0.05) | 0.94(0.06) | 0.02(0.02) | 0.05(0.01) | 0.14(0.02) |
| 100 | 5 | 0.67 | 0.97(0.03) | 0.96(0.03) | 0.02(0.01) | 0.04(0) | 0.11(0.01) |
| 200 | 5 | 0.48 | 0.98(0.01) | 0.97(0.02) | 0.02(0.01) | 0.03(0) | 0.08(0.01) |
| 30 | 10 | 0.64 | 0.94(0.04) | 0.93(0.05) | 0.02(0.01) | 0.04(0.01) | 0.16(0.02) |
| 50 | 10 | 0.71 | 0.97(0.03) | 0.96(0.03) | 0.02(0.01) | 0.03(0) | 0.13(0.02) |
| 100 | 10 | 0.64 | 0.98(0.01) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.1(0.01) |
| 200 | 10 | 0.39 | 0.99(0.01) | 0.98(0.01) | 0.01(0) | 0.02(0) | 0.08(0.01) |
| 30 | 30 | 0.62 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.02(0) | 0.15(0.02) |
| 50 | 30 | 0.7 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.02(0) | 0.12(0.02) |
| 100 | 30 | 0.56 | 0.99(0) | 0.99(0.01) | 0.01(0) | 0.01(0) | 0.1(0.01) |
| 200 | 30 | 0.32 | 0.99(0) | 0.99(0) | 0.01(0) | 0.01(0) | 0.08(0.01) |

Table D.43

*Summary of Fit Statistics Across Conditions: Model M2, Estimator ULSMV, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.07(0.01) | 0.36(0.06) |
| 50 | 5 | 0.99 | 0.99(0.02) | 0.98(0.03) | 0.01(0.01) | 0.06(0.01) | 0.31(0.06) |
| 100 | 5 | 0.94 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.22(0.05) |
| 200 | 5 | 0.93 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.15(0.03) |
| 30 | 10 | 1 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.26(0.04) |
| 50 | 10 | 0.98 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0) | 0.2(0.04) |
| 100 | 10 | 0.94 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.14(0.02) |
| 200 | 10 | 0.87 | 1(0) | 1(0) | 0.01(0) | 0.02(0) | 0.1(0.01) |
| 30 | 30 | 0.99 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.16(0.02) |
| 50 | 30 | 0.96 | 1(0) | 1(0) | 0.01(0) | 0.02(0) | 0.12(0.02) |
| 100 | 30 | 0.84 | 1(0) | 1(0) | 0.01(0) | 0.01(0) | 0.09(0.01) |
| 200 | 30 | 0.62 | 1(0) | 1(0) | 0.01(0) | 0.01(0) | 0.07(0.01) |

Table D.44

*Summary of Fit Statistics Across Conditions: Model M2, Estimator ULSMV, ICC$_O$ 0.1 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.96(0.06) | 0.95(0.07) | 0.01(0.01) | 0.08(0.01) | 0.15(0.04) |
| 50 | 5 | 0.79 | 0.93(0.06) | 0.91(0.07) | 0.02(0.01) | 0.07(0.01) | 0.13(0.03) |
| 100 | 5 | 0.25 | 0.92(0.05) | 0.9(0.06) | 0.03(0.01) | 0.05(0.01) | 0.11(0.02) |
| 200 | 5 | 0 | 0.9(0.04) | 0.88(0.04) | 0.03(0.01) | 0.04(0.01) | 0.1(0.01) |
| 30 | 10 | 0.87 | 0.93(0.08) | 0.91(0.1) | 0.02(0.01) | 0.06(0.01) | 0.13(0.03) |
| 50 | 10 | 0.34 | 0.89(0.06) | 0.87(0.07) | 0.03(0.01) | 0.05(0.01) | 0.11(0.02) |
| 100 | 10 | 0 | 0.87(0.05) | 0.84(0.06) | 0.03(0.01) | 0.04(0.01) | 0.1(0.02) |
| 200 | 10 | 0 | 0.85(0.04) | 0.82(0.05) | 0.03(0) | 0.03(0) | 0.09(0.01) |
| 30 | 30 | 0.88 | 0.95(0.05) | 0.94(0.06) | 0.01(0.01) | 0.04(0.01) | 0.11(0.02) |
| 50 | 30 | 0.12 | 0.87(0.06) | 0.85(0.07) | 0.02(0) | 0.03(0.01) | 0.1(0.02) |
| 100 | 30 | 0 | 0.81(0.06) | 0.77(0.07) | 0.03(0) | 0.03(0.01) | 0.09(0.01) |
| 200 | 30 | 0 | 0.77(0.05) | 0.73(0.06) | 0.03(0) | 0.03(0) | 0.09(0.01) |

Table D.45

*Summary of Fit Statistics Across Conditions: Model M2, Estimator ULSMV, ICC$_O$ 0.3 and ICC$_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.96(0.06) | 0.95(0.07) | 0.01(0.01) | 0.08(0.01) | 0.21(0.04) |
| 50 | 5 | 0.97 | 0.97(0.04) | 0.97(0.04) | 0.01(0.01) | 0.06(0.01) | 0.16(0.02) |
| 100 | 5 | 0.97 | 0.99(0.02) | 0.98(0.02) | 0.01(0.01) | 0.04(0.01) | 0.11(0.01) |
| 200 | 5 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 30 | 10 | 0.99 | 0.98(0.03) | 0.97(0.04) | 0.01(0.01) | 0.05(0.01) | 0.16(0.02) |
| 50 | 10 | 0.98 | 0.98(0.02) | 0.98(0.02) | 0.01(0.01) | 0.04(0) | 0.12(0.01) |
| 100 | 10 | 0.95 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 200 | 10 | 0.92 | 1(0) | 1(0.01) | 0.01(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.03(0) | 0.13(0.01) |
| 50 | 30 | 0.99 | 1(0.01) | 0.99(0.01) | 0(0) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 0.96 | 1(0.01) | 1(0.01) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 0.87 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.46

*Summary of Fit Statistics Across Conditions: Model M2, Estimator ULSMV,* $\text{ICC}_O$ *0.3 and* $\text{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.98 | 0.93(0.08) | 0.92(0.09) | 0.02(0.01) | 0.09(0.01) | 0.14(0.03) |
| 50 | 5 | 0.84 | 0.93(0.06) | 0.92(0.08) | 0.02(0.01) | 0.07(0.01) | 0.11(0.02) |
| 100 | 5 | 0.46 | 0.93(0.05) | 0.91(0.06) | 0.02(0.01) | 0.06(0.01) | 0.09(0.01) |
| 200 | 5 | 0.05 | 0.92(0.03) | 0.9(0.04) | 0.03(0.01) | 0.04(0.01) | 0.07(0.01) |
| 30 | 10 | 0.96 | 0.95(0.07) | 0.93(0.08) | 0.01(0.01) | 0.07(0.01) | 0.12(0.02) |
| 50 | 10 | 0.72 | 0.93(0.06) | 0.91(0.07) | 0.02(0.01) | 0.05(0.01) | 0.09(0.02) |
| 100 | 10 | 0.17 | 0.91(0.05) | 0.89(0.06) | 0.02(0.01) | 0.04(0.01) | 0.08(0.01) |
| 200 | 10 | 0 | 0.9(0.03) | 0.89(0.04) | 0.02(0) | 0.04(0.01) | 0.07(0.01) |
| 30 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.05(0.01) | 0.1(0.02) |
| 50 | 30 | 0.99 | 0.99(0.03) | 0.98(0.03) | 0(0) | 0.04(0.01) | 0.08(0.01) |
| 100 | 30 | 0.26 | 0.93(0.04) | 0.92(0.04) | 0.01(0) | 0.03(0.01) | 0.07(0.01) |
| 200 | 30 | 0 | 0.9(0.03) | 0.88(0.03) | 0.02(0) | 0.03(0.01) | 0.06(0.01) |

Table D.47

*Summary of Fit Statistics Across Conditions: Model M2, Estimator ULSMV,* $\text{ICC}_O$ *0.5 and* $\text{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.96 | 0.9(0.13) | 0.88(0.16) | 0.01(0.01) | 0.09(0.02) | 0.15(0.02) |
| 50 | 5 | 0.97 | 0.95(0.07) | 0.94(0.08) | 0.01(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 5 | 0.97 | 0.97(0.03) | 0.97(0.04) | 0.01(0.01) | 0.04(0.01) | 0.08(0.01) |
| 200 | 5 | 0.94 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 1 | 0.97(0.06) | 0.97(0.07) | 0(0.01) | 0.06(0.01) | 0.13(0.02) |
| 50 | 10 | 0.99 | 0.97(0.04) | 0.96(0.05) | 0.01(0.01) | 0.04(0.01) | 0.1(0.01) |
| 100 | 10 | 0.95 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.03(0) | 0.07(0.01) |
| 200 | 10 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0.01) | 0.12(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.09(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 0.99 | 1(0.01) | 1(0.01) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.48

*Summary of Fit Statistics Across Conditions: Model M2, Estimator ULSMV, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.91(0.11) | 0.89(0.13) | 0.01(0.01) | 0.11(0.02) | 0.12(0.02) |
| 50 | 5 | 0.95 | 0.93(0.08) | 0.92(0.1) | 0.01(0.01) | 0.08(0.01) | 0.09(0.01) |
| 100 | 5 | 0.74 | 0.93(0.05) | 0.92(0.06) | 0.02(0.01) | 0.06(0.01) | 0.07(0.01) |
| 200 | 5 | 0.33 | 0.93(0.04) | 0.92(0.05) | 0.02(0.01) | 0.05(0.01) | 0.06(0.01) |
| 30 | 10 | 1 | 0.98(0.05) | 0.98(0.06) | 0(0.01) | 0.08(0.02) | 0.11(0.02) |
| 50 | 10 | 0.96 | 0.96(0.06) | 0.95(0.07) | 0.01(0.01) | 0.06(0.01) | 0.09(0.01) |
| 100 | 10 | 0.71 | 0.94(0.05) | 0.93(0.06) | 0.01(0.01) | 0.05(0.01) | 0.06(0.01) |
| 200 | 10 | 0.19 | 0.93(0.03) | 0.92(0.04) | 0.02(0) | 0.04(0.01) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.06(0.02) | 0.1(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.05(0.01) | 0.08(0.01) |
| 100 | 30 | 1 | 1(0.01) | 1(0.01) | 0(0) | 0.04(0.01) | 0.06(0.01) |
| 200 | 30 | 0.65 | 0.97(0.03) | 0.96(0.03) | 0.01(0) | 0.03(0.01) | 0.05(0.01) |

Table D.49

*Summary of Fit Statistics Across Conditions: Model M2, Estimator WLSMV, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.98(0.03) | 0.98(0.04) | 0.01(0.01) | 0.07(0.01) | 0.35(0.06) |
| 50 | 5 | 1 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.06(0.01) | 0.3(0.06) |
| 100 | 5 | 0.93 | 0.99(0.02) | 0.98(0.02) | 0.01(0.01) | 0.04(0.01) | 0.22(0.04) |
| 200 | 5 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.15(0.03) |
| 30 | 10 | 1 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.25(0.04) |
| 50 | 10 | 0.96 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0) | 0.19(0.03) |
| 100 | 10 | 0.93 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.13(0.02) |
| 200 | 10 | 0.83 | 1(0) | 0.99(0.01) | 0.01(0) | 0.02(0) | 0.1(0.01) |
| 30 | 30 | 0.98 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.16(0.02) |
| 50 | 30 | 0.95 | 1(0) | 0.99(0.01) | 0.01(0) | 0.02(0) | 0.12(0.02) |
| 100 | 30 | 0.79 | 1(0) | 1(0) | 0.01(0) | 0.01(0) | 0.09(0.01) |
| 200 | 30 | 0.49 | 1(0) | 1(0) | 0.01(0) | 0.01(0) | 0.07(0.01) |

Table D.50

*Summary of Fit Statistics Across Conditions: Model M2, Estimator WLSMV, ICC$_O$ 0.1 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.97(0.03) | 0.97(0.04) | 0.01(0.01) | 0.07(0.01) | 0.19(0.05) |
| 50 | 5 | 0.74 | 0.95(0.04) | 0.94(0.05) | 0.02(0.01) | 0.06(0.01) | 0.15(0.02) |
| 100 | 5 | 0.43 | 0.95(0.03) | 0.94(0.03) | 0.02(0.01) | 0.05(0.01) | 0.12(0.02) |
| 200 | 5 | 0.01 | 0.94(0.02) | 0.93(0.02) | 0.03(0.01) | 0.03(0) | 0.11(0.01) |
| 30 | 10 | 0.97 | 0.97(0.02) | 0.97(0.03) | 0.02(0.01) | 0.05(0.01) | 0.15(0.03) |
| 50 | 10 | 0.7 | 0.97(0.02) | 0.96(0.03) | 0.02(0.01) | 0.04(0.01) | 0.13(0.02) |
| 100 | 10 | 0.05 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.03(0) | 0.11(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.93(0.02) | 0.03(0) | 0.02(0) | 0.1(0.01) |
| 30 | 30 | 0.97 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.13(0.02) |
| 50 | 30 | 0.6 | 0.98(0.01) | 0.98(0.01) | 0.01(0) | 0.02(0) | 0.12(0.01) |
| 100 | 30 | 0 | 0.97(0.01) | 0.97(0.01) | 0.02(0) | 0.02(0) | 0.11(0.01) |
| 200 | 30 | 0 | 0.97(0.01) | 0.96(0.01) | 0.02(0) | 0.01(0) | 0.11(0.01) |

Table D.51

*Summary of Fit Statistics Across Conditions: Model M2, Estimator WLSMV, ICC$_O$ 0.3 and ICC$_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.97 | 0.96(0.04) | 0.96(0.05) | 0.02(0.01) | 0.08(0.01) | 0.21(0.04) |
| 50 | 5 | 0.95 | 0.98(0.03) | 0.97(0.04) | 0.01(0.01) | 0.06(0.01) | 0.15(0.02) |
| 100 | 5 | 0.97 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.11(0.01) |
| 200 | 5 | 0.92 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 30 | 10 | 0.97 | 0.98(0.02) | 0.98(0.03) | 0.01(0.01) | 0.05(0.01) | 0.16(0.02) |
| 50 | 10 | 0.98 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0) | 0.12(0.01) |
| 100 | 10 | 0.94 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 200 | 10 | 0.92 | 1(0) | 1(0) | 0.01(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.13(0.01) |
| 50 | 30 | 0.99 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 0.96 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 0.91 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0.01) |

Table D.52

*Summary of Fit Statistics Across Conditions: Model M2, Estimator WLSMV,* ICC$_O$ *0.3 and* ICC$_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.94 | 0.95(0.05) | 0.94(0.06) | 0.02(0.01) | 0.08(0.01) | 0.16(0.03) |
| 50 | 5 | 0.89 | 0.96(0.03) | 0.95(0.04) | 0.02(0.01) | 0.06(0.01) | 0.13(0.02) |
| 100 | 5 | 0.61 | 0.96(0.02) | 0.96(0.03) | 0.02(0.01) | 0.04(0.01) | 0.1(0.01) |
| 200 | 5 | 0.14 | 0.96(0.02) | 0.96(0.02) | 0.02(0.01) | 0.03(0) | 0.08(0.01) |
| 30 | 10 | 0.97 | 0.98(0.02) | 0.97(0.03) | 0.01(0.01) | 0.05(0.01) | 0.14(0.02) |
| 50 | 10 | 0.85 | 0.98(0.02) | 0.97(0.02) | 0.02(0.01) | 0.04(0.01) | 0.11(0.02) |
| 100 | 10 | 0.34 | 0.97(0.01) | 0.97(0.02) | 0.02(0.01) | 0.03(0) | 0.09(0.01) |
| 200 | 10 | 0.01 | 0.97(0.01) | 0.97(0.01) | 0.02(0) | 0.02(0) | 0.08(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.12(0.02) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 0.52 | 0.99(0) | 0.99(0.01) | 0.01(0) | 0.02(0) | 0.09(0.01) |
| 200 | 30 | 0 | 0.99(0) | 0.99(0) | 0.01(0) | 0.01(0) | 0.08(0.01) |

Table D.53

*Summary of Fit Statistics Across Conditions: Model M2, Estimator WLSMV,* ICC$_O$ *0.5 and* ICC$_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.95 | 0.96(0.05) | 0.95(0.06) | 0.02(0.01) | 0.08(0.01) | 0.16(0.02) |
| 50 | 5 | 0.97 | 0.98(0.03) | 0.97(0.03) | 0.01(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 5 | 0.97 | 0.99(0.01) | 0.99(0.02) | 0.01(0.01) | 0.04(0.01) | 0.08(0.01) |
| 200 | 5 | 0.96 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 0.99 | 0.99(0.02) | 0.99(0.02) | 0.01(0.01) | 0.05(0.01) | 0.13(0.01) |
| 50 | 10 | 0.99 | 0.99(0.01) | 0.99(0.01) | 0.01(0.01) | 0.04(0.01) | 0.1(0.01) |
| 100 | 10 | 0.96 | 1(0.01) | 0.99(0.01) | 0.01(0.01) | 0.03(0) | 0.07(0.01) |
| 200 | 10 | 0.94 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.12(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.09(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.07(0.01) |
| 200 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.05(0) |

Table D.54

*Summary of Fit Statistics Across Conditions: Model M2, Estimator WLSMV, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.98 | 0.96(0.04) | 0.95(0.05) | 0.02(0.01) | 0.08(0.01) | 0.15(0.02) |
| 50 | 5 | 0.95 | 0.97(0.03) | 0.96(0.04) | 0.02(0.01) | 0.06(0.01) | 0.11(0.01) |
| 100 | 5 | 0.77 | 0.97(0.02) | 0.97(0.03) | 0.02(0.01) | 0.05(0.01) | 0.08(0.01) |
| 200 | 5 | 0.42 | 0.98(0.01) | 0.97(0.02) | 0.02(0.01) | 0.03(0) | 0.06(0.01) |
| 30 | 10 | 0.99 | 0.99(0.02) | 0.99(0.02) | 0(0.01) | 0.05(0.01) | 0.13(0.02) |
| 50 | 10 | 0.97 | 0.99(0.01) | 0.98(0.02) | 0.01(0.01) | 0.04(0.01) | 0.1(0.01) |
| 100 | 10 | 0.77 | 0.99(0.01) | 0.98(0.01) | 0.01(0.01) | 0.03(0) | 0.08(0.01) |
| 200 | 10 | 0.27 | 0.99(0.01) | 0.98(0.01) | 0.01(0) | 0.02(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.03(0) | 0.12(0.02) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.1(0.01) |
| 100 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.02(0) | 0.08(0.01) |
| 200 | 30 | 0.73 | 1(0) | 1(0) | 0(0) | 0.01(0) | 0.06(0.01) |

Table D.55

*Summary of Fit Statistics Across Conditions: Model M12, Estimator MLR, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.1 | 0.77(0.14) | 0.73(0.16) | 0.07(0.02) | 0.08(0.04) | 0.32(0.06) |
| 50 | 5 | 0.1 | 0.86(0.06) | 0.83(0.07) | 0.05(0.01) | 0.06(0.01) | 0.28(0.04) |
| 100 | 5 | 0.01 | 0.89(0.04) | 0.87(0.04) | 0.04(0.01) | 0.05(0.01) | 0.22(0.03) |
| 200 | 5 | 0 | 0.91(0.02) | 0.9(0.03) | 0.04(0) | 0.05(0.01) | 0.17(0.02) |
| 30 | 10 | 0.03 | 0.86(0.05) | 0.83(0.06) | 0.05(0.01) | 0.06(0.01) | 0.26(0.04) |
| 50 | 10 | 0.01 | 0.89(0.04) | 0.87(0.04) | 0.04(0.01) | 0.05(0.01) | 0.22(0.03) |
| 100 | 10 | 0 | 0.91(0.02) | 0.9(0.03) | 0.04(0) | 0.05(0) | 0.16(0.02) |
| 200 | 10 | 0 | 0.92(0.01) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.12(0.02) |
| 30 | 30 | 0 | 0.9(0.03) | 0.88(0.03) | 0.04(0.01) | 0.05(0) | 0.19(0.03) |
| 50 | 30 | 0 | 0.91(0.02) | 0.89(0.02) | 0.04(0) | 0.04(0) | 0.15(0.02) |
| 100 | 30 | 0 | 0.91(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.11(0.01) |
| 200 | 30 | 0 | 0.91(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.08(0.01) |

Table D.56

*Summary of Fit Statistics Across Conditions: Model M12, Estimator MLR, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.03 | 0.83(0.06) | 0.8(0.07) | 0.07(0.02) | 0.08(0.01) | 0.2(0.04) |
| 50 | 5 | 0.01 | 0.87(0.04) | 0.85(0.05) | 0.06(0.01) | 0.07(0.01) | 0.17(0.03) |
| 100 | 5 | 0 | 0.89(0.03) | 0.88(0.03) | 0.05(0.01) | 0.06(0.01) | 0.14(0.02) |
| 200 | 5 | 0 | 0.9(0.02) | 0.88(0.02) | 0.05(0) | 0.05(0.01) | 0.13(0.02) |
| 30 | 10 | 0.02 | 0.86(0.05) | 0.84(0.05) | 0.06(0.01) | 0.06(0.01) | 0.17(0.04) |
| 50 | 10 | 0 | 0.89(0.03) | 0.87(0.03) | 0.05(0.01) | 0.05(0.01) | 0.15(0.03) |
| 100 | 10 | 0 | 0.9(0.02) | 0.88(0.02) | 0.05(0) | 0.05(0.01) | 0.13(0.02) |
| 200 | 10 | 0 | 0.9(0.01) | 0.89(0.01) | 0.04(0) | 0.05(0) | 0.12(0.01) |
| 30 | 30 | 0 | 0.9(0.02) | 0.88(0.03) | 0.04(0.01) | 0.05(0) | 0.16(0.04) |
| 50 | 30 | 0 | 0.91(0.02) | 0.89(0.02) | 0.04(0) | 0.04(0) | 0.14(0.02) |
| 100 | 30 | 0 | 0.91(0.01) | 0.89(0.01) | 0.04(0) | 0.04(0) | 0.13(0.02) |
| 200 | 30 | 0 | 0.91(0.01) | 0.89(0.01) | 0.04(0) | 0.04(0) | 0.12(0.01) |

Table D.57

*Summary of Fit Statistics Across Conditions: Model M12, Estimator MLR, $\mathrm{ICC}_O$ 0.3 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.28 | 0.8(0.13) | 0.76(0.15) | 0.05(0.02) | 0.07(0.01) | 0.23(0.03) |
| 50 | 5 | 0.35 | 0.88(0.07) | 0.86(0.08) | 0.04(0.01) | 0.06(0.01) | 0.19(0.02) |
| 100 | 5 | 0.14 | 0.91(0.05) | 0.89(0.05) | 0.03(0.01) | 0.05(0.01) | 0.14(0.02) |
| 200 | 5 | 0 | 0.91(0.03) | 0.9(0.03) | 0.03(0.01) | 0.05(0.01) | 0.1(0.01) |
| 30 | 10 | 0.13 | 0.87(0.06) | 0.84(0.07) | 0.04(0.01) | 0.06(0.01) | 0.2(0.02) |
| 50 | 10 | 0.06 | 0.9(0.04) | 0.88(0.05) | 0.04(0.01) | 0.05(0.01) | 0.15(0.02) |
| 100 | 10 | 0 | 0.91(0.02) | 0.89(0.03) | 0.03(0.01) | 0.05(0.01) | 0.11(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.08(0.01) |
| 30 | 30 | 0 | 0.9(0.03) | 0.88(0.03) | 0.04(0.01) | 0.05(0.01) | 0.16(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.13(0.01) |
| 100 | 30 | 0 | 0.91(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.09(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.06(0.01) |

Table D.58

*Summary of Fit Statistics Across Conditions: Model M12, Estimator MLR, $\text{ICC}_O$ 0.3 and $\text{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.17 | 0.83(0.09) | 0.8(0.1) | 0.06(0.02) | 0.08(0.01) | 0.18(0.03) |
| 50 | 5 | 0.15 | 0.88(0.05) | 0.86(0.06) | 0.05(0.01) | 0.07(0.01) | 0.15(0.02) |
| 100 | 5 | 0.01 | 0.9(0.03) | 0.89(0.04) | 0.04(0.01) | 0.05(0.01) | 0.12(0.02) |
| 200 | 5 | 0 | 0.91(0.02) | 0.89(0.02) | 0.04(0) | 0.05(0) | 0.1(0.01) |
| 30 | 10 | 0.06 | 0.87(0.05) | 0.85(0.06) | 0.05(0.01) | 0.06(0.01) | 0.17(0.03) |
| 50 | 10 | 0.01 | 0.9(0.03) | 0.88(0.04) | 0.04(0.01) | 0.05(0.01) | 0.13(0.02) |
| 100 | 10 | 0 | 0.91(0.02) | 0.89(0.03) | 0.04(0) | 0.05(0) | 0.1(0.01) |
| 200 | 10 | 0 | 0.91(0.01) | 0.9(0.02) | 0.04(0) | 0.04(0) | 0.09(0.01) |
| 30 | 30 | 0 | 0.9(0.02) | 0.89(0.03) | 0.04(0.01) | 0.05(0) | 0.15(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.9(0.02) | 0.04(0) | 0.04(0) | 0.12(0.02) |
| 100 | 30 | 0 | 0.91(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.1(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.09(0.01) |

Table D.59

*Summary of Fit Statistics Across Conditions: Model M12, Estimator MLR, $\text{ICC}_O$ 0.5 and $\text{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.38 | 0.81(0.1) | 0.78(0.12) | 0.05(0.02) | 0.07(0.01) | 0.19(0.02) |
| 50 | 5 | 0.39 | 0.88(0.07) | 0.86(0.08) | 0.03(0.01) | 0.06(0.01) | 0.15(0.02) |
| 100 | 5 | 0.21 | 0.91(0.04) | 0.89(0.05) | 0.03(0.01) | 0.05(0.01) | 0.11(0.01) |
| 200 | 5 | 0.01 | 0.91(0.03) | 0.9(0.03) | 0.03(0) | 0.05(0.01) | 0.07(0.01) |
| 30 | 10 | 0.26 | 0.88(0.06) | 0.86(0.07) | 0.04(0.01) | 0.06(0.01) | 0.17(0.02) |
| 50 | 10 | 0.09 | 0.9(0.04) | 0.88(0.04) | 0.03(0.01) | 0.05(0.01) | 0.13(0.01) |
| 100 | 10 | 0 | 0.91(0.03) | 0.9(0.03) | 0.03(0) | 0.04(0) | 0.09(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.06(0.01) |
| 30 | 30 | 0 | 0.9(0.02) | 0.88(0.03) | 0.04(0) | 0.05(0) | 0.16(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.12(0.01) |
| 100 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.06(0.01) |

Table D.60

*Summary of Fit Statistics Across Conditions: Model M12, Estimator MLR, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.3 | 0.82(0.09) | 0.79(0.11) | 0.05(0.02) | 0.07(0.01) | 0.18(0.03) |
| 50 | 5 | 0.36 | 0.89(0.07) | 0.87(0.08) | 0.04(0.01) | 0.06(0.01) | 0.13(0.02) |
| 100 | 5 | 0.07 | 0.9(0.04) | 0.89(0.04) | 0.03(0.01) | 0.05(0.01) | 0.1(0.01) |
| 200 | 5 | 0 | 0.91(0.02) | 0.9(0.03) | 0.03(0) | 0.05(0.01) | 0.08(0.01) |
| 30 | 10 | 0.18 | 0.88(0.06) | 0.85(0.07) | 0.04(0.01) | 0.06(0.01) | 0.16(0.02) |
| 50 | 10 | 0.07 | 0.9(0.04) | 0.89(0.04) | 0.03(0.01) | 0.05(0.01) | 0.12(0.02) |
| 100 | 10 | 0 | 0.91(0.02) | 0.9(0.03) | 0.03(0) | 0.04(0) | 0.09(0.01) |
| 200 | 10 | 0 | 0.92(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.07(0.01) |
| 30 | 30 | 0 | 0.9(0.03) | 0.89(0.03) | 0.03(0.01) | 0.04(0) | 0.15(0.02) |
| 50 | 30 | 0 | 0.91(0.02) | 0.9(0.02) | 0.03(0) | 0.04(0) | 0.12(0.02) |
| 100 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.09(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.9(0.01) | 0.03(0) | 0.04(0) | 0.07(0.01) |

Table D.61

*Summary of Fit Statistics Across Conditions: Model M12, Estimator ULSMV, $\mathrm{ICC}_O$ 0.1 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 1 | 0.95(0.05) | 0.94(0.05) | 0.02(0.01) | 0.09(0.01) | 0.35(0.06) |
| 50 | 5 | 0.61 | 0.93(0.05) | 0.92(0.06) | 0.03(0.01) | 0.07(0.01) | 0.32(0.06) |
| 100 | 5 | 0.09 | 0.92(0.03) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.22(0.05) |
| 200 | 5 | 0 | 0.93(0.02) | 0.91(0.03) | 0.04(0.01) | 0.05(0.01) | 0.15(0.03) |
| 30 | 10 | 0.43 | 0.93(0.04) | 0.91(0.04) | 0.03(0.01) | 0.07(0.01) | 0.26(0.05) |
| 50 | 10 | 0.05 | 0.93(0.03) | 0.91(0.03) | 0.03(0.01) | 0.06(0.01) | 0.2(0.04) |
| 100 | 10 | 0 | 0.92(0.02) | 0.91(0.02) | 0.04(0.01) | 0.05(0.01) | 0.13(0.02) |
| 200 | 10 | 0 | 0.92(0.01) | 0.91(0.02) | 0.04(0) | 0.05(0) | 0.09(0.01) |
| 30 | 30 | 0.01 | 0.93(0.02) | 0.92(0.02) | 0.03(0.01) | 0.05(0.01) | 0.16(0.02) |
| 50 | 30 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.12(0.02) |
| 100 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.04(0) | 0.05(0) | 0.09(0.01) |
| 200 | 30 | 0 | 0.92(0.01) | 0.91(0.01) | 0.04(0) | 0.05(0) | 0.06(0.01) |

Table D.62

*Summary of Fit Statistics Across Conditions: Model M12, Estimator ULSMV, ICC$_O$ 0.1 and ICC$_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.82 | 0.9(0.12) | 0.88(0.14) | 0.02(0.02) | 0.09(0.01) | 0.16(0.04) |
| 50 | 5 | 0.44 | 0.87(0.07) | 0.85(0.08) | 0.03(0.01) | 0.08(0.01) | 0.14(0.03) |
| 100 | 5 | 0.06 | 0.87(0.04) | 0.85(0.05) | 0.04(0.01) | 0.07(0.01) | 0.12(0.02) |
| 200 | 5 | 0 | 0.86(0.04) | 0.83(0.05) | 0.04(0.01) | 0.06(0.01) | 0.11(0.02) |
| 30 | 10 | 0.47 | 0.86(0.08) | 0.83(0.1) | 0.03(0.01) | 0.07(0.01) | 0.15(0.04) |
| 50 | 10 | 0.14 | 0.85(0.07) | 0.82(0.08) | 0.03(0.01) | 0.06(0.01) | 0.12(0.03) |
| 100 | 10 | 0 | 0.82(0.05) | 0.79(0.06) | 0.04(0.01) | 0.06(0.01) | 0.1(0.02) |
| 200 | 10 | 0 | 0.77(0.05) | 0.73(0.05) | 0.04(0) | 0.05(0.01) | 0.11(0.02) |
| 30 | 30 | 0.36 | 0.88(0.05) | 0.86(0.06) | 0.02(0.01) | 0.06(0.01) | 0.12(0.03) |
| 50 | 30 | 0 | 0.81(0.05) | 0.78(0.06) | 0.03(0) | 0.06(0) | 0.11(0.02) |
| 100 | 30 | 0 | 0.75(0.03) | 0.7(0.04) | 0.03(0) | 0.05(0) | 0.09(0.02) |

Table D.63

*Summary of Fit Statistics Across Conditions: Model M12, Estimator ULSMV, ICC$_O$ 0.3 and ICC$_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.95 | 0.92(0.08) | 0.9(0.09) | 0.02(0.01) | 0.09(0.01) | 0.21(0.04) |
| 50 | 5 | 0.76 | 0.92(0.06) | 0.91(0.07) | 0.02(0.01) | 0.07(0.01) | 0.16(0.02) |
| 100 | 5 | 0.3 | 0.93(0.04) | 0.92(0.04) | 0.03(0.01) | 0.06(0.01) | 0.11(0.01) |
| 200 | 5 | 0.02 | 0.93(0.02) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.08(0.01) |
| 30 | 10 | 0.84 | 0.92(0.05) | 0.91(0.06) | 0.02(0.01) | 0.07(0.01) | 0.16(0.02) |
| 50 | 10 | 0.38 | 0.93(0.03) | 0.91(0.04) | 0.02(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 10 | 0.02 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.08(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.93(0.02) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 0.93 | 0.97(0.03) | 0.96(0.03) | 0.01(0.01) | 0.06(0.01) | 0.13(0.01) |
| 50 | 30 | 0.31 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.05(0) | 0.1(0.01) |
| 100 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.02(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.02(0) | 0.05(0) | 0.05(0.01) |

Table D.64

*Summary of Fit Statistics Across Conditions: Model M12, Estimator ULSMV, $\mathrm{ICC}_O$ 0.3 and $\mathrm{ICC}_L$ 0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.93 | 0.9(0.1) | 0.88(0.12) | 0.02(0.01) | 0.11(0.01) | 0.14(0.03) |
| 50 | 5 | 0.66 | 0.89(0.08) | 0.87(0.09) | 0.03(0.01) | 0.09(0.01) | 0.11(0.02) |
| 100 | 5 | 0.17 | 0.89(0.05) | 0.87(0.06) | 0.03(0.01) | 0.07(0.01) | 0.09(0.01) |
| 200 | 5 | 0 | 0.89(0.03) | 0.87(0.04) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 30 | 10 | 0.89 | 0.91(0.08) | 0.89(0.1) | 0.02(0.01) | 0.08(0.01) | 0.12(0.02) |
| 50 | 10 | 0.48 | 0.89(0.06) | 0.87(0.07) | 0.02(0.01) | 0.07(0.01) | 0.1(0.02) |
| 100 | 10 | 0.02 | 0.88(0.04) | 0.85(0.05) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0 | 0.87(0.03) | 0.84(0.04) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0.02) | 1(0.02) | 0(0) | 0.06(0.01) | 0.1(0.02) |
| 50 | 30 | 0.93 | 0.97(0.04) | 0.96(0.04) | 0.01(0) | 0.06(0.01) | 0.08(0.01) |
| 100 | 30 | 0.01 | 0.9(0.03) | 0.88(0.04) | 0.02(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.86(0.03) | 0.84(0.03) | 0.02(0) | 0.05(0) | 0.06(0.01) |

Table D.65

*Summary of Fit Statistics Across Conditions: Model M12, Estimator ULSMV, $\mathrm{ICC}_O$ 0.5 and $\mathrm{ICC}_L$ 0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.95 | 0.87(0.14) | 0.84(0.17) | 0.02(0.01) | 0.1(0.02) | 0.15(0.02) |
| 50 | 5 | 0.94 | 0.9(0.09) | 0.89(0.1) | 0.02(0.01) | 0.08(0.01) | 0.12(0.01) |
| 100 | 5 | 0.77 | 0.93(0.05) | 0.92(0.06) | 0.02(0.01) | 0.06(0.01) | 0.08(0.01) |
| 200 | 5 | 0.36 | 0.94(0.03) | 0.93(0.03) | 0.02(0) | 0.06(0.01) | 0.06(0.01) |
| 30 | 10 | 0.99 | 0.95(0.07) | 0.94(0.09) | 0.01(0.01) | 0.07(0.01) | 0.13(0.01) |
| 50 | 10 | 0.94 | 0.93(0.06) | 0.92(0.07) | 0.01(0.01) | 0.06(0.01) | 0.1(0.01) |
| 100 | 10 | 0.66 | 0.94(0.04) | 0.93(0.04) | 0.01(0.01) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0.13 | 0.94(0.02) | 0.93(0.02) | 0.02(0) | 0.05(0) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.06(0.01) | 0.12(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.05(0) | 0.09(0.01) |
| 100 | 30 | 0.99 | 0.98(0.02) | 0.98(0.02) | 0(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0.22 | 0.96(0.01) | 0.95(0.02) | 0.01(0) | 0.05(0) | 0.05(0) |

Table D.66

*Summary of Fit Statistics Across Conditions: Model M12, Estimator ULSMV,* $\mathrm{ICC}_O$ *0.5 and* $\mathrm{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.99 | 0.89(0.12) | 0.87(0.14) | 0.02(0.01) | 0.12(0.02) | 0.12(0.02) |
| 50 | 5 | 0.93 | 0.91(0.09) | 0.89(0.1) | 0.02(0.01) | 0.09(0.01) | 0.09(0.01) |
| 100 | 5 | 0.57 | 0.9(0.06) | 0.89(0.07) | 0.02(0.01) | 0.08(0.01) | 0.07(0.01) |
| 200 | 5 | 0.11 | 0.91(0.04) | 0.89(0.04) | 0.02(0) | 0.06(0.01) | 0.05(0.01) |
| 30 | 10 | 0.99 | 0.97(0.07) | 0.96(0.08) | 0(0.01) | 0.09(0.01) | 0.11(0.02) |
| 50 | 10 | 0.94 | 0.93(0.07) | 0.92(0.08) | 0.01(0.01) | 0.08(0.01) | 0.08(0.01) |
| 100 | 10 | 0.48 | 0.91(0.05) | 0.9(0.06) | 0.02(0.01) | 0.06(0.01) | 0.06(0.01) |
| 200 | 10 | 0.02 | 0.91(0.03) | 0.89(0.04) | 0.02(0) | 0.06(0.01) | 0.05(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.08(0.01) | 0.1(0.01) |
| 50 | 30 | 1 | 1(0) | 1(0) | 0(0) | 0.07(0.01) | 0.08(0.01) |
| 100 | 30 | 1 | 0.99(0.02) | 0.99(0.02) | 0(0) | 0.06(0.01) | 0.06(0.01) |
| 200 | 30 | 0.32 | 0.94(0.03) | 0.93(0.03) | 0.01(0) | 0.05(0) | 0.05(0.01) |

Table D.67

*Summary of Fit Statistics Across Conditions: Model M12, Estimator WLSMV,* $\mathrm{ICC}_O$ *0.1 and* $\mathrm{ICC}_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.94 | 0.94(0.05) | 0.93(0.06) | 0.02(0.01) | 0.09(0.01) | 0.35(0.06) |
| 50 | 5 | 0.68 | 0.93(0.05) | 0.92(0.06) | 0.03(0.01) | 0.07(0.01) | 0.3(0.05) |
| 100 | 5 | 0.2 | 0.93(0.03) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.22(0.04) |
| 200 | 5 | 0 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.05(0.01) | 0.15(0.03) |
| 30 | 10 | 0.62 | 0.93(0.04) | 0.92(0.05) | 0.03(0.01) | 0.07(0.01) | 0.25(0.05) |
| 50 | 10 | 0.15 | 0.93(0.03) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.19(0.03) |
| 100 | 10 | 0 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.05(0.01) | 0.13(0.02) |
| 200 | 10 | 0 | 0.93(0.01) | 0.92(0.02) | 0.03(0) | 0.05(0) | 0.09(0.01) |
| 30 | 30 | 0 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.05(0.01) | 0.16(0.02) |
| 50 | 30 | 0 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.05(0) | 0.12(0.01) |
| 100 | 30 | 0 | 0.93(0.01) | 0.92(0.01) | 0.03(0) | 0.05(0) | 0.09(0.01) |
| 200 | 30 | 0 | 0.93(0.01) | 0.91(0.01) | 0.03(0) | 0.05(0) | 0.06(0.01) |

Table D.68

*Summary of Fit Statistics Across Conditions: Model M12, Estimator WLSMV,* ICC$_O$ *0.1 and* ICC$_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.79 | 0.91(0.09) | 0.89(0.1) | 0.03(0.02) | 0.09(0.01) | 0.2(0.06) |
| 50 | 5 | 0.22 | 0.89(0.04) | 0.87(0.05) | 0.04(0.01) | 0.08(0.01) | 0.16(0.03) |
| 100 | 5 | 0.02 | 0.9(0.04) | 0.88(0.04) | 0.04(0.01) | 0.06(0.01) | 0.12(0.02) |
| 200 | 5 | 0 | 0.88(0.03) | 0.86(0.03) | 0.04(0) | 0.06(0.01) | 0.11(0.02) |
| 30 | 10 | 0.34 | 0.91(0.06) | 0.89(0.07) | 0.03(0.01) | 0.07(0.01) | 0.16(0.04) |
| 50 | 10 | 0.03 | 0.89(0.04) | 0.87(0.05) | 0.04(0.01) | 0.06(0.01) | 0.12(0.02) |
| 100 | 10 | 0 | 0.87(0.04) | 0.85(0.04) | 0.04(0.01) | 0.06(0.01) | 0.1(0.02) |
| 200 | 10 | 0 | 0.87(0.02) | 0.84(0.03) | 0.04(0) | 0.05(0) | 0.1(0.02) |
| 30 | 30 | 0.03 | 0.92(0.03) | 0.91(0.03) | 0.03(0.01) | 0.06(0.01) | 0.13(0.02) |
| 50 | 30 | 0 | 0.9(0.02) | 0.88(0.03) | 0.03(0) | 0.05(0) | 0.11(0.02) |
| 100 | 30 | 0 | 0.87(0.02) | 0.84(0.03) | 0.04(0) | 0.05(0) | 0.1(0.01) |
| 200 | 30 | 0 | 0.86(0.02) | 0.83(0.02) | 0.04(0) | 0.05(0) | 0.09(0.01) |

Table D.69

*Summary of Fit Statistics Across Conditions: Model M12, Estimator WLSMV,* ICC$_O$ *0.3 and* ICC$_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.9 | 0.91(0.07) | 0.9(0.08) | 0.03(0.01) | 0.09(0.01) | 0.21(0.04) |
| 50 | 5 | 0.63 | 0.93(0.05) | 0.91(0.06) | 0.03(0.01) | 0.07(0.01) | 0.16(0.02) |
| 100 | 5 | 0.21 | 0.93(0.04) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.11(0.01) |
| 200 | 5 | 0.01 | 0.93(0.02) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.08(0.01) |
| 30 | 10 | 0.57 | 0.93(0.04) | 0.91(0.05) | 0.03(0.01) | 0.07(0.01) | 0.15(0.02) |
| 50 | 10 | 0.16 | 0.93(0.03) | 0.91(0.04) | 0.03(0.01) | 0.06(0.01) | 0.12(0.01) |
| 100 | 10 | 0.01 | 0.93(0.02) | 0.92(0.02) | 0.03(0) | 0.06(0.01) | 0.08(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.92(0.02) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 0.14 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.06(0.01) | 0.13(0.01) |
| 50 | 30 | 0 | 0.94(0.01) | 0.93(0.02) | 0.03(0) | 0.05(0) | 0.1(0.01) |
| 100 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.03(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.92(0.01) | 0.03(0) | 0.05(0) | 0.05(0.01) |

Table D.70

*Summary of Fit Statistics Across Conditions: Model M12, Estimator WLSMV,* ICC$_O$ *0.3 and* ICC$_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.81 | 0.91(0.07) | 0.89(0.08) | 0.03(0.01) | 0.09(0.01) | 0.16(0.03) |
| 50 | 5 | 0.51 | 0.91(0.05) | 0.89(0.06) | 0.03(0.01) | 0.08(0.01) | 0.13(0.02) |
| 100 | 5 | 0.06 | 0.91(0.03) | 0.9(0.04) | 0.03(0.01) | 0.06(0.01) | 0.09(0.01) |
| 200 | 5 | 0 | 0.91(0.02) | 0.9(0.03) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 30 | 10 | 0.53 | 0.92(0.04) | 0.91(0.05) | 0.03(0.01) | 0.07(0.01) | 0.14(0.02) |
| 50 | 10 | 0.11 | 0.92(0.03) | 0.9(0.04) | 0.03(0.01) | 0.06(0.01) | 0.11(0.02) |
| 100 | 10 | 0 | 0.91(0.03) | 0.9(0.03) | 0.03(0) | 0.06(0.01) | 0.08(0.01) |
| 200 | 10 | 0 | 0.91(0.02) | 0.89(0.02) | 0.04(0) | 0.05(0) | 0.07(0.01) |
| 30 | 30 | 0.74 | 0.97(0.02) | 0.96(0.02) | 0.01(0.01) | 0.06(0.01) | 0.12(0.02) |
| 50 | 30 | 0.04 | 0.95(0.02) | 0.94(0.02) | 0.02(0) | 0.05(0.01) | 0.09(0.01) |
| 100 | 30 | 0 | 0.93(0.01) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.08(0.01) |
| 200 | 30 | 0 | 0.91(0.01) | 0.9(0.01) | 0.03(0) | 0.05(0) | 0.06(0.01) |

Table D.71

*Summary of Fit Statistics Across Conditions: Model M12, Estimator WLSMV,* ICC$_O$ *0.5 and* ICC$_L$ *0.1*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.8 | 0.91(0.07) | 0.89(0.08) | 0.03(0.02) | 0.09(0.01) | 0.15(0.02) |
| 50 | 5 | 0.7 | 0.92(0.05) | 0.91(0.06) | 0.03(0.01) | 0.08(0.01) | 0.12(0.01) |
| 100 | 5 | 0.27 | 0.93(0.03) | 0.92(0.04) | 0.03(0.01) | 0.06(0.01) | 0.08(0.01) |
| 200 | 5 | 0 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.06(0.01) |
| 30 | 10 | 0.83 | 0.95(0.04) | 0.94(0.04) | 0.02(0.01) | 0.07(0.01) | 0.13(0.01) |
| 50 | 10 | 0.29 | 0.94(0.03) | 0.92(0.03) | 0.03(0.01) | 0.06(0.01) | 0.1(0.01) |
| 100 | 10 | 0 | 0.94(0.02) | 0.93(0.02) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0 | 0.94(0.01) | 0.93(0.02) | 0.03(0) | 0.05(0) | 0.05(0.01) |
| 30 | 30 | 0.98 | 0.99(0.02) | 0.99(0.02) | 0(0.01) | 0.06(0.01) | 0.12(0.01) |
| 50 | 30 | 0.41 | 0.97(0.01) | 0.96(0.02) | 0.01(0) | 0.05(0.01) | 0.09(0.01) |
| 100 | 30 | 0 | 0.95(0.01) | 0.94(0.01) | 0.02(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.03(0) | 0.05(0) | 0.05(0) |

Table D.72

*Summary of Fit Statistics Across Conditions: Model M12, Estimator WLSMV,* $\mathrm{ICC}_O$ *0.5 and* $\mathrm{ICC}_L$ *0.5*

| $N_2$ | $N_1$ | $\chi^2$ | CFI | TLI | RMSEA | SRMRW | SRMRB |
|---|---|---|---|---|---|---|---|
| 30 | 5 | 0.87 | 0.92(0.06) | 0.9(0.07) | 0.03(0.01) | 0.09(0.01) | 0.14(0.02) |
| 50 | 5 | 0.65 | 0.92(0.05) | 0.91(0.06) | 0.03(0.01) | 0.08(0.01) | 0.11(0.01) |
| 100 | 5 | 0.15 | 0.92(0.03) | 0.91(0.04) | 0.03(0.01) | 0.07(0.01) | 0.08(0.01) |
| 200 | 5 | 0 | 0.92(0.02) | 0.91(0.03) | 0.03(0) | 0.06(0.01) | 0.06(0.01) |
| 30 | 10 | 0.87 | 0.95(0.04) | 0.95(0.05) | 0.02(0.01) | 0.07(0.01) | 0.13(0.02) |
| 50 | 10 | 0.35 | 0.94(0.03) | 0.93(0.04) | 0.03(0.01) | 0.06(0.01) | 0.1(0.01) |
| 100 | 10 | 0.01 | 0.93(0.02) | 0.92(0.03) | 0.03(0) | 0.06(0.01) | 0.07(0.01) |
| 200 | 10 | 0 | 0.93(0.02) | 0.91(0.02) | 0.03(0) | 0.05(0) | 0.06(0.01) |
| 30 | 30 | 1 | 1(0) | 1(0.01) | 0(0) | 0.06(0.01) | 0.12(0.01) |
| 50 | 30 | 0.85 | 0.98(0.02) | 0.98(0.02) | 0.01(0) | 0.06(0.01) | 0.09(0.01) |
| 100 | 30 | 0 | 0.95(0.01) | 0.94(0.01) | 0.02(0) | 0.05(0) | 0.07(0.01) |
| 200 | 30 | 0 | 0.94(0.01) | 0.93(0.01) | 0.03(0) | 0.05(0) | 0.05(0.01) |

# APPENDIX E

## Additional Figures of the Distribution of Fit Indices

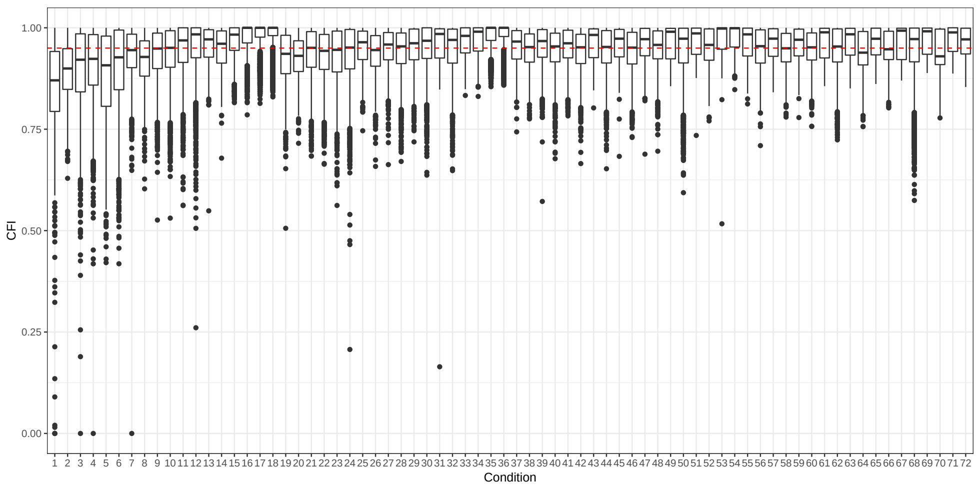



*Figure E.1.* Distribution of CFI across all conditions

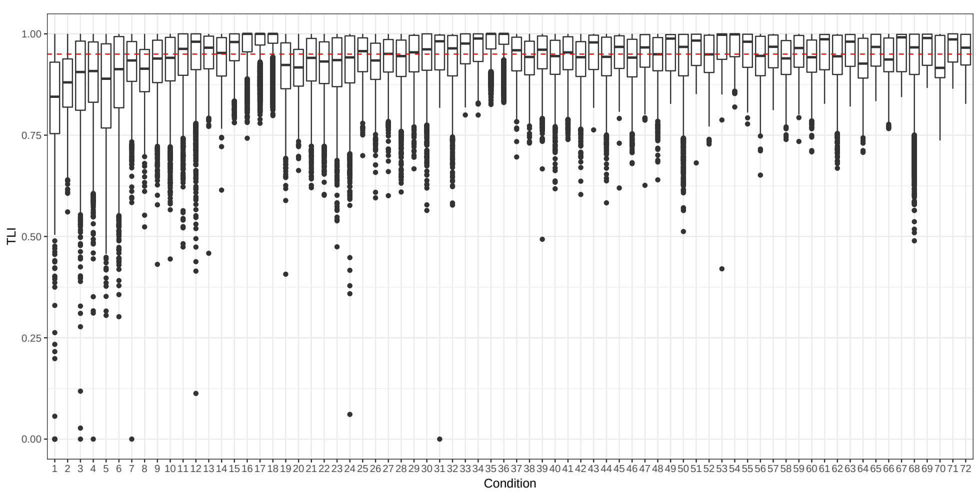



*Figure E.2.* Distribution of TLI across all conditions

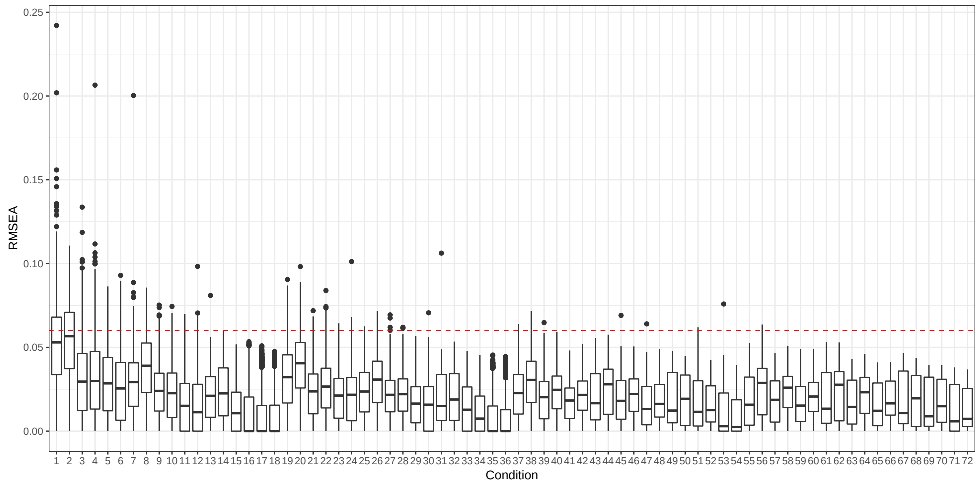



*Figure E.3.* Distribution of RMSEA across all conditions

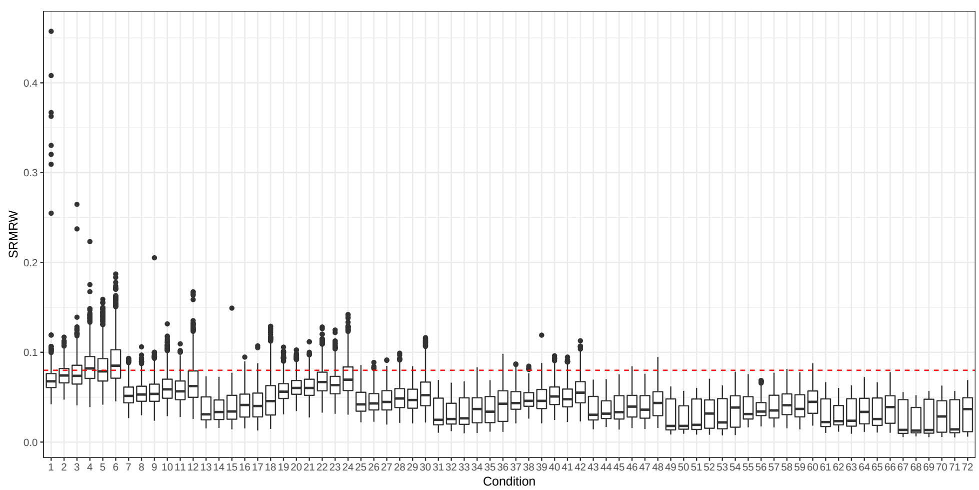



*Figure E.4.* Distribution of SRMRW across all conditions

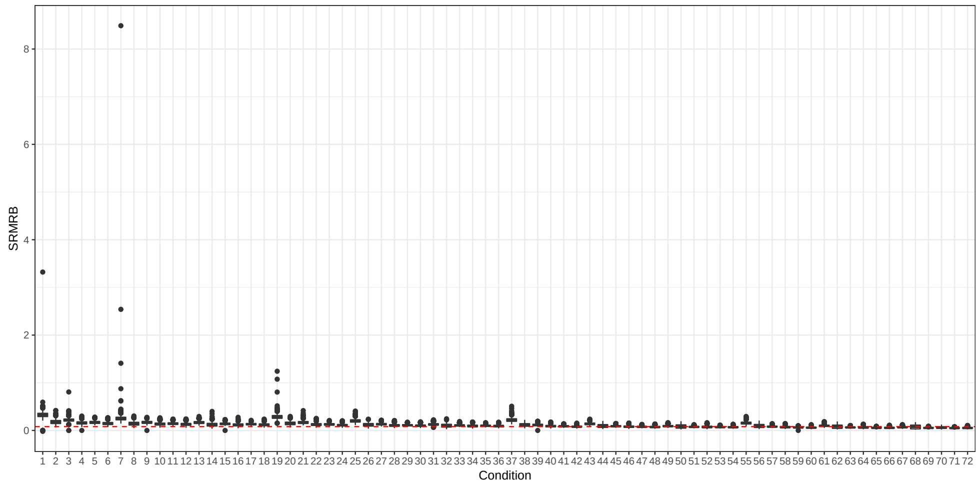



*Figure E.5.* Distribution of SRMRB across all conditions

*Note.* Dashed (red) line represents the Hu & Benter (1999) commonly reported cutoff for SRMR at .08. The excessively large values observed in Conditions 1 and 7 occured when number of groups (Ng=30) with 5 and 10 units per group respectively and low ICCs. The max value observed (8.49) for apparently admissible solution occurred under WLSMV, Model M1, and according to CFI, TLI, and RMSEA, this model fits perfectly, and the $\chi^2$ test of goodness of fit was not reject. Highest value observed for Model C was 2.54, in the same condition 7.

REFERENCES


Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society. Series A (General)*, *149*(1), 1. doi: 10.2307/2981882

Asparouhov, T., & Muthén, B. (2007). Computationally efficient estimation of multilevel high-dimensional latent variable models. In *Joint statistical meeting: Section on statistics in epidemiology.* Retrieved from https://www.statmodel.com/download/JSM2007000746.pdf

Babakus, E., Ferguson, C. E., & Joreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research*, *24*(2), 222. doi: 10.2307/3151512

Bandalos, D. L. (2014). Relative performance of categorical diagonally weighted least squares and robust maximum likelihood estimation. *Structural Equation Modeling*, *21*(1), 102–116. doi: 10.1080/10705511.2014.859510

Bell, B. A., Morgan, G. B., Schoeneberger, J. A., Kromrey, J. D., & Ferron, J. M. (2014). How low can you go?: An investigation of the influence of sample size and model complexity on point and interval estimates in two-level linear models. *Methodology*, *10*(1), 1–11. doi: 10.1027/1614-2241/a000062

Benson, J., & Nasser, F. (1998). On the use of factor analysis as a research tool. *Journal of Vocational Education Research*, *23*(1), 13.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. doi: 10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonnett, D. (1980). Significance tests and goodness of fit. *Psychological Bulletin*, *88*(3), 588–606.

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

Bollen, K. A., & Barb, K. H. (1981). Pearson's R and coarsely categorized measures. *American Sociological Review*, *46*(2), 232–239. doi: 10.2307/2094981

Bollen, K. A., Bauer, D. J., Christ, S. L., & Edwards, M. C. (2010). Overview of structural equation models and recent extensions. In *Statistics in the social sciences: Current methodological developments* (pp. 37–79). doi: 10.1002/9780470583333.ch2

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* Guilford Press.

Bryk, A. S., & Driscoll, M. E. (1988). *An empirical investigation of school as a community* (Tech. Rep.). Madison, WI: University of Wisconsin Research Center on Effective Secondary Schools.

DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment*, *23*(3), 225–241. doi: 10.1177/073428290502300303

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling*, *21*(3), 425–438. doi: 10.1080/10705511.2014.915373

Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, *47*(2), 309–326. doi: 10.1111/j.2044-8317.1994.tb01039.x

Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *Leadership Quarterly*, *16*(1), 149–167. doi: 10.1016/j.leaqua.2004.09.009

Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 439–492). Charlotte, NC: Information Age Publishing.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological methods*, *9*(4), 466–491. doi: 10.1037/1082-989X.9.4.466

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS sstimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 625–641. doi: 10.1080/10705510903203573

Goldstein, H., & McDonald, R. P. (1988). A general model for the analysis of multilevel data. *Psychometrika*, *53*(4), 455–467. doi: 0033-3123/88/0900-9022500.75/0

Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(4), 621–638. doi: 10.1080/10705511.2017.1402334

Harwell, M. R. (1992). Summarizing Monte Carlo results in methodological research. *Journal of Educational Statistics*, *17*(4), 297–313. doi: 10.3102/10769986017004297

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus.* (Third ed.). New York, NY: Routledge.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management*, *23*(6), 723–744. doi: 10.1177/014920639702300602

Hox, J. J., & Maas, C. J. (2001). The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples. *Structural Equation Modeling*, *8*(2), 157–174. doi: 10.1207/S15328007SEM0802_1

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.

Hsu, H. Y. (2009). *Testing the effectiveness of various commonly used fit indices for detecting misspecifications in multilevel structure equation models* (Unpublished doctoral dissertation). Texas A&M University.

Hsu, H. Y., Kwok, O., Lin, H., & Acosta, S. (2015). Detecting isspecified multilevel structural equation models with common fit indices: A monte carlo study. *Multivariate Behavioral Research*, *50*(2), 197–215. doi: 10.1080/00273171.2014.977429

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. doi: 1082-989X/98/J3.00

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. doi: 10.1080/10705519909540118

Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review*, *48*(3), 398–407.

Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, *32*(4), 443–482. doi: 10.1007/BF02289658

Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, *34*(2), 183–202.

Jöreskog, K. G. (2005). *Structural equation modeling with ordinal variables* (Tech. Rep.). Lincolnwood, IL. doi: 10.1214/lnms/1215463803

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by maximum likelihood and least squares methods.* Chicago, IL: National Educational Resources.

Jöreskog, K. G., & Sörbom, D. (2015). *LISREL 9.20 for Windows.* Skokie, IL: Scientific Software International, Inc.

Kaplan, D. (2000). Multilevel structural equation modeling. In *Structural equation modeling* (First ed., pp. 136–140). Thousand Oaks, CA: SAGE Publications Inc.

Kaplan, D. (2009). *Structural equation modeling* (Second ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Kim, E. S., Dedrick, R. F., Cao, C., & Ferron, J. M. (2016). Multilevel factor fnalysis: Reporting guidelines and a review of reporting practices. *Multivariate Behavioral Research*, *51*(6), 881–898. doi: 10.1080/00273171.2016.1228042

Kline, R. B. (2016). *Principles and practice of structural equation modeling.* Guilford Press.

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. doi: 10.3758/s13428-015-0619-7

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. doi: 10.1027/1614-2241.1.3.86

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, N.J.: Routledge.

McDonald, R. P., & Goldstein, H. (1989). Balanced versus unbalanced designs for linear structural relations in two-level data. *British Journal of Mathematical and Statistical Psychology*, *42*(2), 215–232. doi: 10.1111/j.2044-8317.1989.tb00911.x

Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, *49*(1), 115–132. doi: 10.1007/BF02294210

Muthén, B. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika*, *54*(4), 557–585.

Muthén, B. (1990). *Means and covariance structure analysis of hierarchical data* (Tech. Rep.). Los Angeles, CA. Retrieved from https://escholarship.org/uc/item/1vp6w4sr

Muthén, B. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, *22*(3), 376–398. doi: 10.1177/0049124194022003006

Muthén, B., du Toit, S., & Spisic, D. (1997). *Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes* (Tech. Rep.).

Muthén, B., & 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*(2), 171–189. doi: 10.1111/j.2044-8317.1985.tb00832.x

Muthen, B., & Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of the model. *British Journal of Mathematical and Statistical Psychology*, *45*(1), 19–30. doi: 10.1111/j.2044-8317.1992.tb00975.x

Muthén, B., & Satorra, A. (1989). Multilevel aspects of varying parameters in structural models. In R. D. Bock (Ed.), *Mulitlevel analysis of educational data* (pp. 87–99). San Diego: Academic Press Inc.

Muthén, B. O., & Asparouhov, T. (n.d.). *Technical appendix: SRMR* (Tech. Rep.). Muthén & Muthén. Retrieved from http://www.statmodel.com/techappen.shtml

Muthén, L., & Muthén, B. (2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Navruz, B. (2016). *The behaviors of robust weighted least squares estimation techniques for categorical/ordinal data in multilevel CFA models* (Unpublished doctoral dissertation). Texas A&M University.

Nestler, S. (2013). A Monte Carlo study comparing PIV, ULS and DWLS in the estimation of dichotomous confirmatory factor analysis. *British Journal of Mathematical and Statistical Psychology*, *66*(1), 127–143. doi: 10.1111/j.2044-8317.2012.02044.x

Nye, C. D., & Drasgow, F. (2011). Assessing goodness of fit: Simple rules of thumb simply do not work. *Organizational Research Methods*, *14*(3), 548–570. doi: 10.1177/1094428110368562

Padgett, R. N. (2019). *MCFA thesis results summary.* Texas Data Repository Dataverse. doi: 10.18738/T8/VPAH7O

R Core Team. (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Second ed., Vol. 1). SAGE Publications, Inc.

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*(3), 354–373. doi: 10.1037/a0029315

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal*, *32*(3), 583–625. doi: 10.3102/00028312032003583

Ryu, E. (2014). Model fit evaluation in multilevel structural equation models. *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.00081

Ryu, E., & West, S. G. (2009). Level-specific evaluation of model fit in multilevel structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 583–601. doi: 10.1080/10705510903203466

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347–367. doi: 10.1177/1094428107308906

Snijders, T. A. B., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational and Behavioral Statistics*, *18*(3), 237–259. doi: 10.3102/10769986018003237

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling.* Thousand Oaks, CA: SAGE Publications, Inc.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology*, *15*(2), 201. doi: 10.2307/1412107

Stapleton, L. M. (2013). Multilevel structural equation modeling with complex sample data. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 521–562). Charlotte, NC: Information Age Publishing.

Stapleton, L. M., Yang, J. S., & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics*, *41*(5), 481–520. doi: 10.3102/1076998616646200

Steiger, J. H., & Lind, J. C. (1980). Statistically based tests for the number of common factors. In *The annual meeting of the psychometric society* (pp. 424–453). Iowa City, Iowa.

Stochl, J., Jones, P. B., Perez, J., Khandaker, G. M., Böhnke, J. R., & Croudace, T. J. (2016). Effects of ignoring clustered data structure in confirmatory factor analysis of ordered polytomous items: a simulation study based on PANSS. *International Journal of Methods in Psychiatric Research*, *25*(3), 205–219. doi: 10.1002/mpr.1474

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. doi: 10.1007/BF02291170

Van der Leeden, R., & Busing, F. (1994). First iteration versus IGLS/RIGLS estimates in two-level models: A Monte Carlo study with ML3 (Technical Report PRM 94-03). *Department of Psychometrics and Research Methodology, University of Leiden, Leiden, the Netherlands.*

Wu, J., & Kwok, O. (2012). Using SEM to analyze complex survey data: A comparison between design-based single-level and model-based multilevel approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(1), 16–35. doi: 10.1080/10705511.2012.634703

Yang-Wallentin, F., Jöreskog, K. G., & Luo, H. (2010). Confirmatory factor analysis of ordinal variables with misspecified models. *Structural Equation Modeling*, *17*(3), 392–423. doi: 10.1080/10705511.2010.489003