An Equivalence Testing Based Version of the Standardized Root Mean Squared Residual

Nataly Beribisky and Robert Cribbie
Department of Psychology, York University

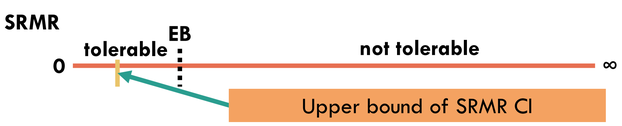
1 INTRODUCTION

- Fit indices are used to describe model fit in structural equation modeling (SEM).
- For instance, the standardized root mean squared residual (SRMR) measures the mean value of residual correlation left after an SEM model has been fit to the data.
- Fit indices were meant to be used as effect sizes but often function as informal check tests (ICTs).
- To disentangle fit indices from ICTs, an inferential test, like an equivalence test (ET), may be introduced to evaluate model fit.
- There are already ETs for certain fit indices but there is not yet one for SRMR.

The present study had two goals:
- Propose variations of SRMR ETs.
- Compare the performance of these tests to one another and to ICTs using a Monte Carlo simulation study.

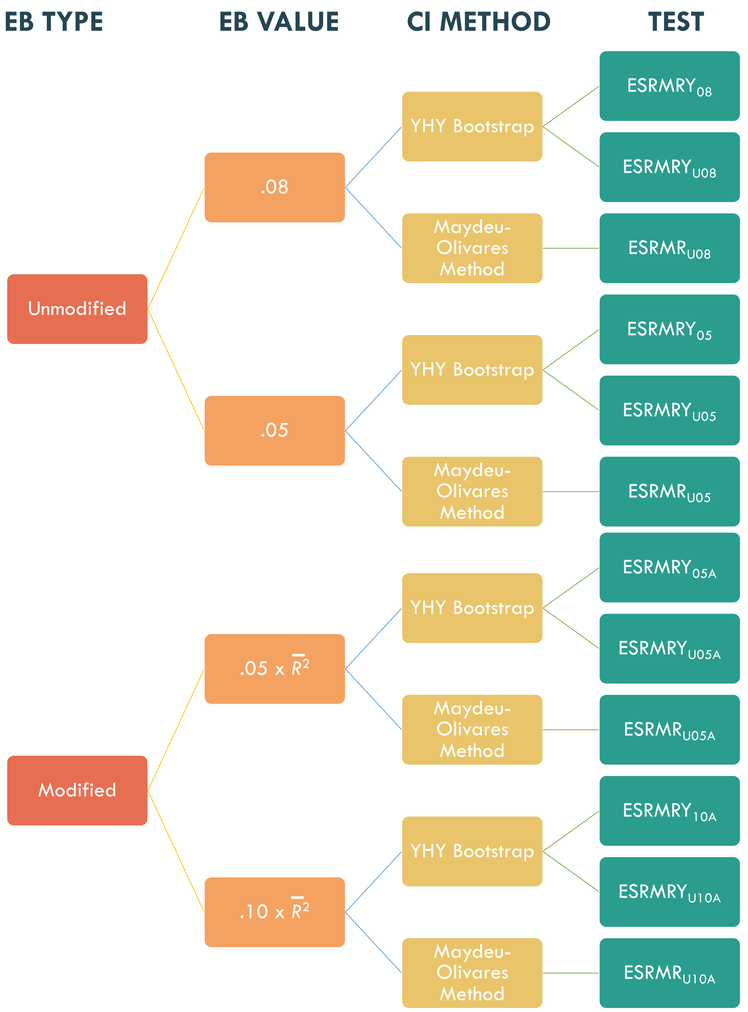
2 DEVELOPING EQUIVALENCE TESTS

- In SEM, equivalence tests compare the misspecification in an identified model to a minimally tolerable size of misspecification.
- This involves comparing a given bound of a fit index's confidence interval (CI) to an equivalence bound (EB; one bound of an equivalence interval).



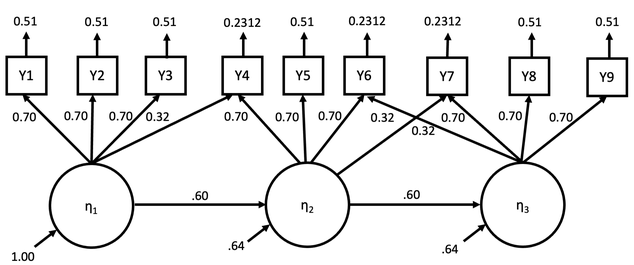
3 PROPOSED EQUIVALENCE TESTS

- EBs were either **unmodified** or **modified:**
  - Unmodified EBs: .05 or .08.
  - Modified EBs: proposed by Shi et al. (2018) and Shi et al. (2022) where .05 or .10 was multiplied by the average R^2 of the observed indicators ($\overline{R}^2$).
- The ET was based on either the **original SRMR** (Bentler, 1995) or the **unbiased SRMR** (Maydeu-Olivares, 2017).
- The confidence interval was either computed via a **YHY bootstrap** (Yuan et al., 2007) or via a method derived by **Maydeu-Olivares** (2017).



4 MONTE CARLO METHOD

- We adapted three population generating models from Chen et al. (2007) for our simulation study.
- The two primary factors manipulated were sample size and model misspecification:
  - N = 50, 75, 100, 200, 400, 800, 1000, and 5000.
  - Misspecification type = (1) non-negligible (SRMR ≥ EB), (2) negligible (SRMR < EB), and (3) perfect.
- Replications = 1000 & bootstraps samples = 500.



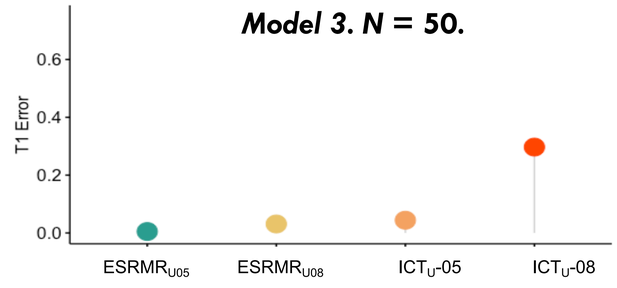
Example population generating model.

5 RESULTS

Non-Negligible Misspecification

- Error rates were generally higher in small sample sizes for all ETs with errors occurring under 5% of the time for all models/sample sizes.
- ICTs tended to have error rates that were similar or higher than the corresponding equivalence tests.
- When misspecification was set at the EB, tests using the Maydeu-Olivares CI had error control within the Bradley (1978) liberal criterion.

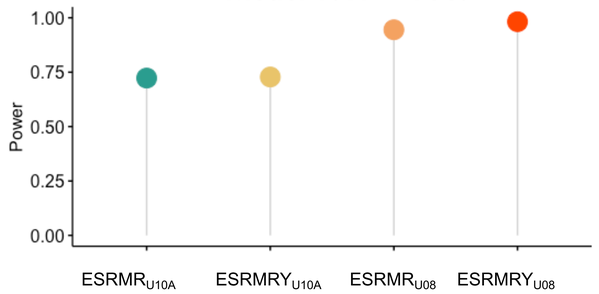
Model 3. N = 50.



Negligible Misspecification

- All ETs reached a power of ~1 by N = 5000.
- $ESRMRY_{U08}$ reached the power ceiling the quickest of all ETs, but $ESRMR_{U08}$ had the highest power at the lowest sample sizes, and was comparable to $ESRMRY_{U08}$ by N = 200.

Model 1. N = 100.



Perfect Fit

- Most ETs reached a power of 1 by N = 1000.
- Again, $ESRMRY_{U08}$ reached the power ceiling the quickest, but $ESRMR_{U08}$ was relatively comparable to $ESRMRY_{U08}$ at all sample sizes and models.

6 DISCUSSION

- We recommend the $ESRMR_{U08}$ ET due to its good combination of power and Type I error control.
- We hope equivalence tests for SRMR assist SEM researchers evaluating model fit.