Examining SEM Trees for Investigating Measurement Invariance Concerning Multiple Violators

2023 Modern Modeling Methods Conference
June 27, 2023

Yuanfang Liu
University of Cincinnati
Mark Lai
University of Southern California

University of CINCINNATI

Outline

- Research Background on Measurement Invariance
- Structural Equation Model (SEM) Trees
- Simulation Design
- Results
- Research Significance

University of CINCINNATI

Measurement Invariance (MI)

Measurement Invariance (a property)

A construct is measured in conceptually similar ways across different groups (Vandenberg & Lance, 2000).

Measurement Invariance

- Under a common factor model, measurement (factorial) invariance is expressed as

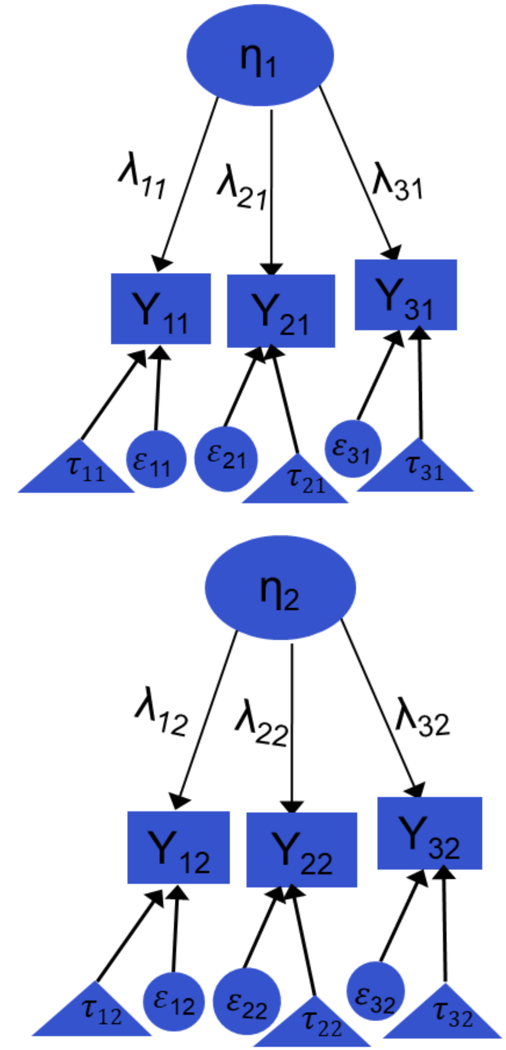
$$\mathbf{Y}_k = \boldsymbol{\tau}_k + \boldsymbol{\Lambda}_k \boldsymbol{\eta}_k + \boldsymbol{\varepsilon}_k$$

$\mathbf{Y}_{\boldsymbol{k}}$, observed item scores vector for *p* items, subpopulation *k*
$\boldsymbol{\tau}_{\boldsymbol{k}}$, measurement intercepts vector
$\boldsymbol{\eta}_{\boldsymbol{k}}$, a vector for *m* common factors
$\boldsymbol{\Lambda}_k$, a $p \times m$ factor loading matrix
$\boldsymbol{\varepsilon}_k$, unique factor score vector

University of CINCINNATI

Four Stages of Factorial Invariance

- Configural invariance
- Metric (Weak) invariance (equal item factor loadings λ)
- Scalar (Strong) invariance (equal item intercepts τ)
- Strict invariance (equal unique factor ε variances and covariances; Meredith, 1993)



University of CINCINNATI

Measurement Invariance Methods

- Multiple-group confirmatory factor analysis (Jöreskog, 1971)
- Multiple indicators, multiple causes (MIMIC; Muthén,1989)
- Alignment optimization (Asparouhov & Muthén, 2014)
- Structural equation model (SEM) trees (Brandmaier et al., 2016)

  Previous application of SEM trees focused on uncovering general data heterogeneity.

University of CINCINNATI

SEM Trees

- **Structural equation modeling**

(causal hypotheses & causal relations)

- **Decision trees**

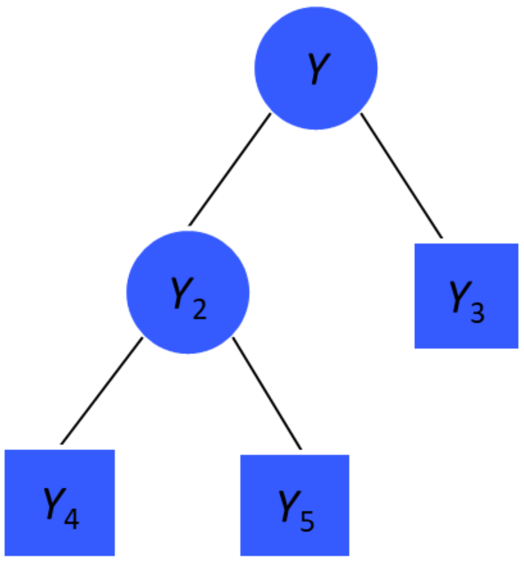
(Partitioning; Breiman et al., 1998)

- **Structural equation model trees** (Brandmaier et al., 2016)

Partition a dataset concerning a model

Observed covariates

Predict parameter estimate differences



Maximum Likelihood of a SEM tree

- -2LL $(T \mid D) = \sum_{d \in D} -2\text{LL}(M(\theta_{\psi(T,d)}) \mid d)$

T tree structure

LL loglikelihood

D a $n \times (p + q)$ data set matrix

n sample size

p number of observed indicators for model/template M

q number of covariates

$\psi(T, d)$ a mapping function between an observation d to a node of T

$\theta_{\psi(T,d)}$ parameter estimates to a node of T

University of CINCINNATI

Tree Growth under LR test

- LR test statistic

$\Lambda = -2[\text{LL}\,(\hat{\theta}_F \mid D_F) - \sum_{i=1}^{k} \text{LL}\,(\hat{\theta}_i \mid D_i)]$

Asymptotically χ^2 distributed

D_F full data set before split

m number of free parameters under the model *M*

$i = 1, \ldots, k$ number of sub data sets

$df = (k-1)m$

SEM Tree Measurement Invariance

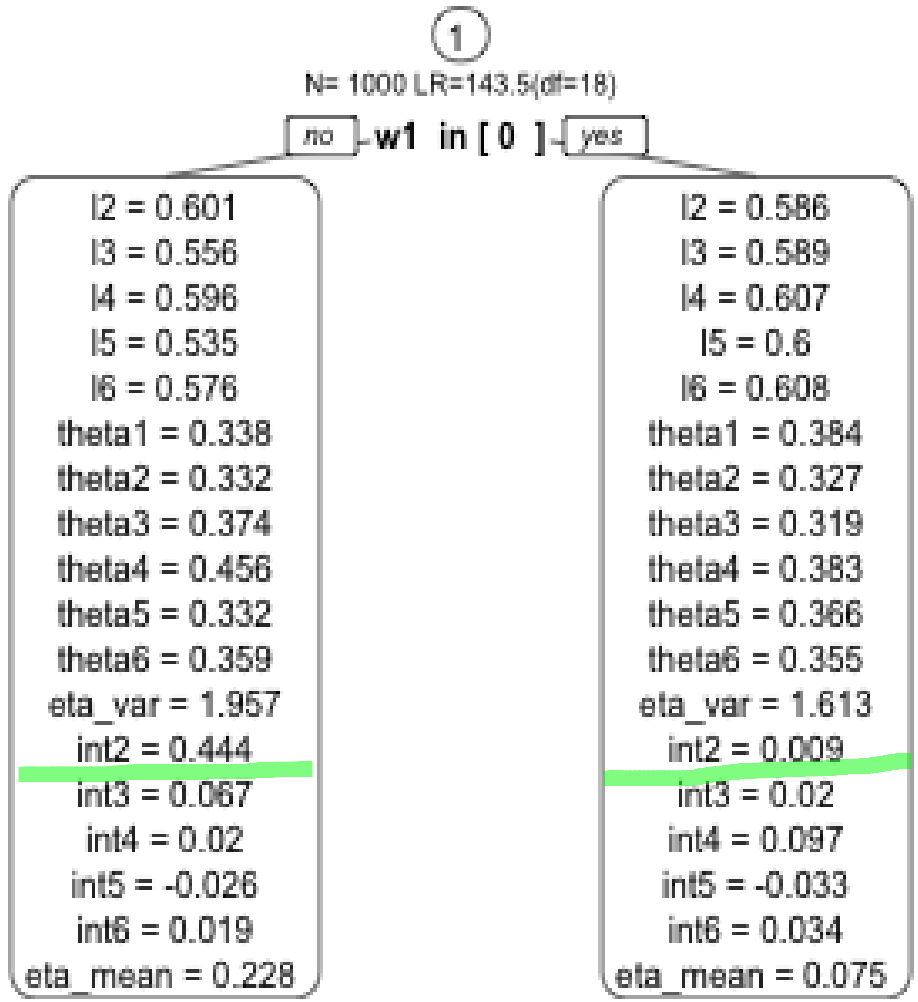
- Local invariance (SEM target parameters are approximately equal across final groups)

  Estimate parameters for each tree node during tree growth
- Global invariance (SEM target parameters are exactly equal across all inner nodes and final groups)

  Estimate parameters once and then fix them
- Number of free parameters estimated under the two constraints are different.

University of CINCINNATI

SEM Tree Example under Intercept Noninvariance (One Replication)



University of CINCINNATI

Simulation Design

Under a one factor model with six items, 250 Replications

- **Sample size**
- **Type of Noninvariance**
- **Number of noninvariant items**
- **Magnitude of noninvariance**
- **Types of Violators**
- **Number of Violators**
- **Relationships between a violator and noninvariant item(s)**

University of CINCINNATI

Analytic Models & Testing Procedure

In the original presentation, we showed the analytic models and testing procedures for identifying noninvariance via likelihood comparisons.

Evaluation Criteria

Primary Criteria

Type I error rates: the proportion of replications that M_2 was incorrectly selected in full invariance conditions.

Statistical power rates: the proportion of replications that M_2 was correctly selected in intercept noninvariance conditions.

University of CINCINNATI

Evaluation Criteria (cont.)

Split rates (SR): the proportion of replications that a covariate served as a group membership for at least one time under the intercept noninvariance model during data split across replications.

Secondary Criteria

Range, Bias, and Root mean square error (RMSE) of loadings and intercepts parameters of the leaves under selected model

University of CINCINNATI

Result

- Type I error rates $\leq .052$, $n \leq 1000$

- Statistical power rates in .40-1.00 (n = 500) and .96-1.00 (n = 1000)

- SR

University of CINCINNATI

Sample Size	SR			
	W_1 (Dichotomous Violator)	W_2 (Continuous Violator)	W_3 (Dichotomous Noise Covariate)	W_4 (Continuous Noise Covariate)
Small and Linear Intercept Noninvariance				
500	.400–.932	.948–1.00	.000–.004	.000–.060
1000	.928–1.00	1.00	.000–.008	.008–.064
Medium and Linear				
500	.964–1.00	1.00	.000–.012	.012–.116
1000	.972–1.00	1.00	.000–.012	.008–.088
Small and Nonlinear				
500	---	.460–1.00	.000–.012	.028–.052
1000	---	.980–1.00	.000–.016	.052–.084
Medium and Nonlinear				
500	---	.904–1.00	.012–.020	.048–.100
1000	---	1.00	.008–.024	.080–**.148**

Result (Cont.)

- a noise covariate for noninvariance & causal indicator of a latent construct

  W_1, SR in .000-.024 (n = 500) and .008-.036 (n = 1000)

  W_2, SR in .048-.100 (n = 500) and .304-**.312** (n = 1000)

- SEM tree had high **SR ≥ .928** (n =1000) for dichotomous violators while keeping **low SR ≤ .024** for **dichotomous noise covariates, meaning that** a dichotomous group membership for tree split was very likely to be a violator contributing to intercept noninvariance.

University of CINCINNATI

Research Significance

- A first study using SEM tree to investigate measurement noninvariance concerning multiple violators

- SEM tree performed well in detecting both linear and nonlinear intercept noninvariance

- An exploratory procedure to identify target parameter differences, which might contribute to theory revision on a related SEM framework, construct development, and item design

University of CINCINNATI

- Thank you!
- liu2yf@mail.uc.edu
- *Note*

  The PDF is simplified, compared to the in-person presentation version.

References

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling*, *21*(4), 495–508.

Brandmaier, A. M., Prindle, J. J., McArdle, J. J., & Lindenberger, U. (2016). Theory-guided exploration with structural equation model forests. *Psychological Methods*, *21*(4), 566–582.

Breiman, Friedman, Olshen, & Stone. (1998). *Classification and regression trees*. CRC Press. (Original work published 1984).

University of CINCINNATI

References (cont.)

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*(4), 409–426.

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

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*(1), 4–70.

University of CINCINNATI