Using a Parametric Boostrap Approach for Nested Model Comparisons in Structural Equation Modeling

Sunthud Pornprasertmanit
Wei Wu
Todd D. Little
University of Kansas

KU CENTER FOR RESEARCH METHODS & DATA ANALYSIS

College of Liberal Arts & Sciences

Outline

- Introduction
- Trivial Misspecification
- Research Questions
- Parametric Bootstrap Approach
- Simulation
- Result and Discussion

KU CENTER FOR
RESEARCH METHODS
& DATA ANALYSIS

College of Liberal Arts & Sciences

Introduction

- Model evaluation
- Chi-square
- Practical fit Indices
  - Can a single cutoff be applied to any models with any sample sizes and with any types of data?
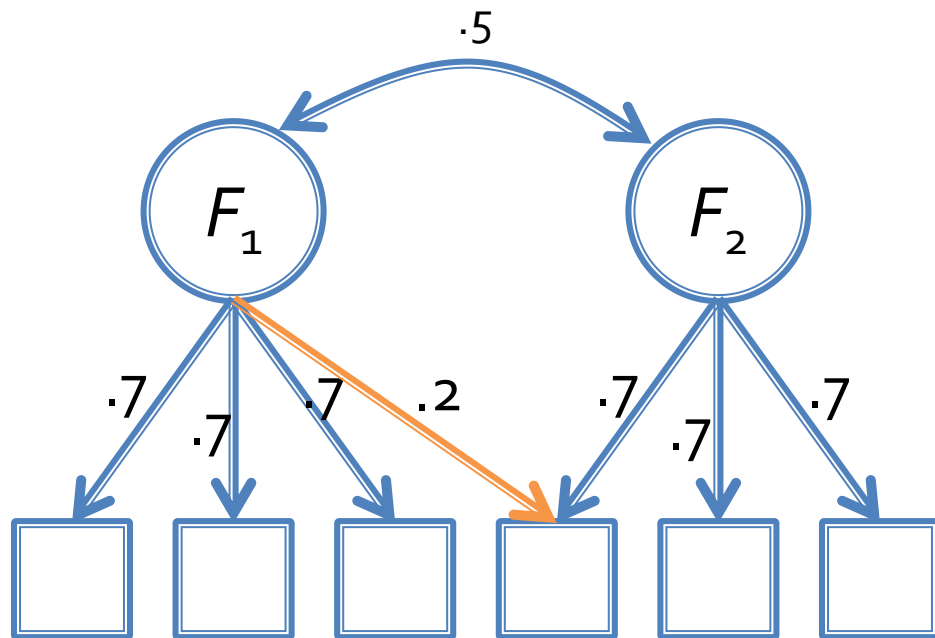
Introduction

- Nested Model Comparison
- Chi-square difference test
- Change in fit indices (e.g., ΔCFI)
  - Results of simulation study → Golden rules
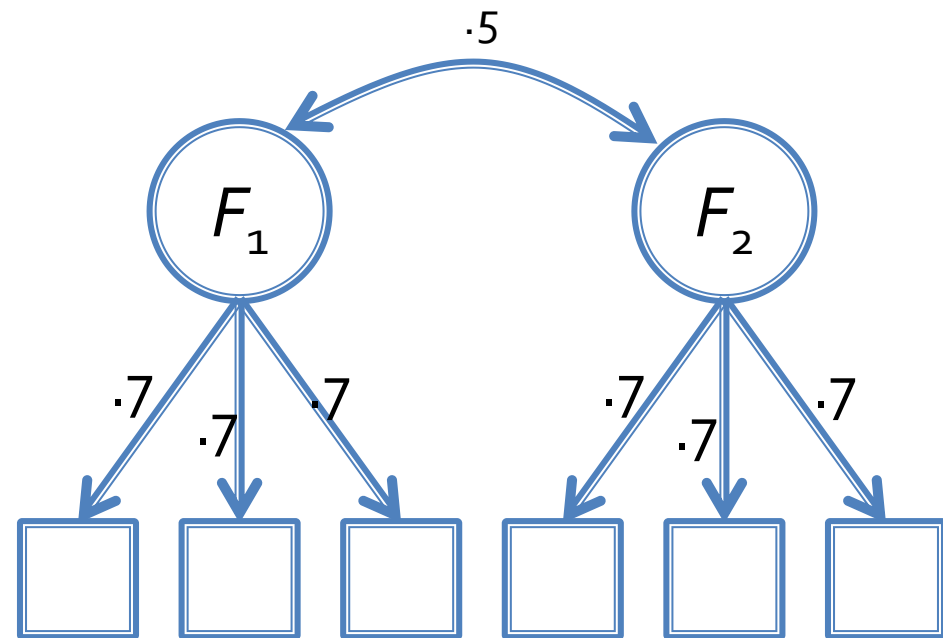 - Creative applications to nongeneralizable situations

Trivial Misspecification

- Hypothesized model approximates the underlying mechanism.

Research Questions

- How can the cutoff be tailored for each hypothesized model?
- How can the golden rule be avoided?
- How can the cutoff be accounted for trivial misspecification? → Test of approximate fit
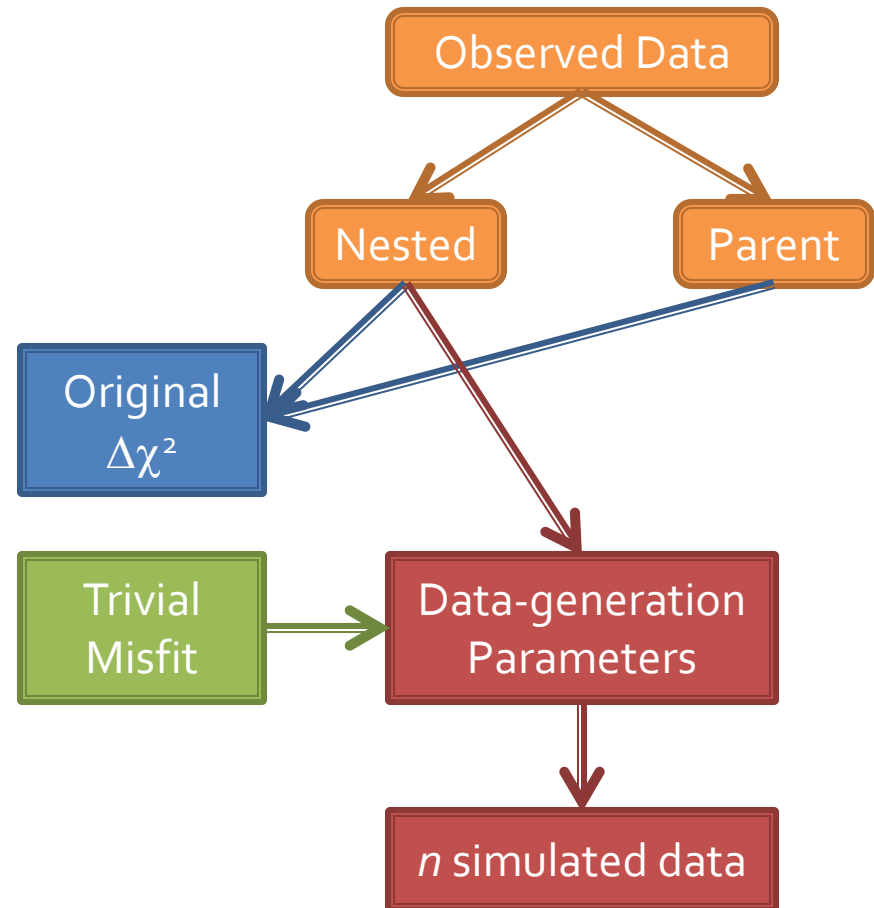
Parametric Bootstrap

- Fit the nested and parent models to the observed data
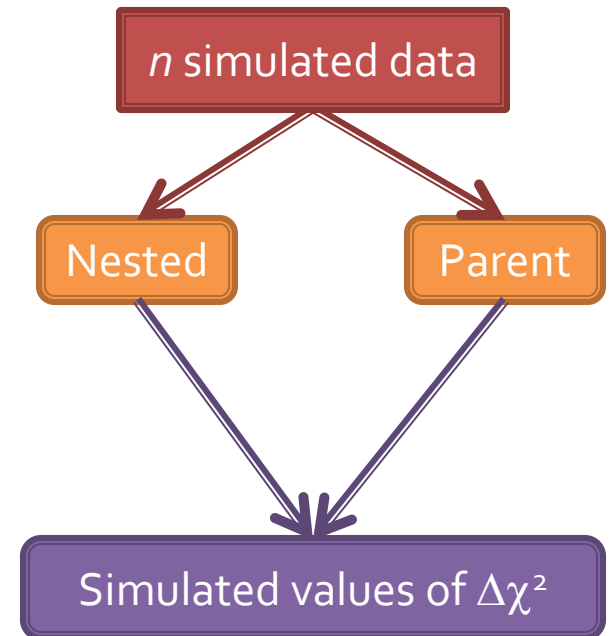- Save the difference in a fit index

Observed Data → Nested, Parent → Original $\Delta\chi^2$
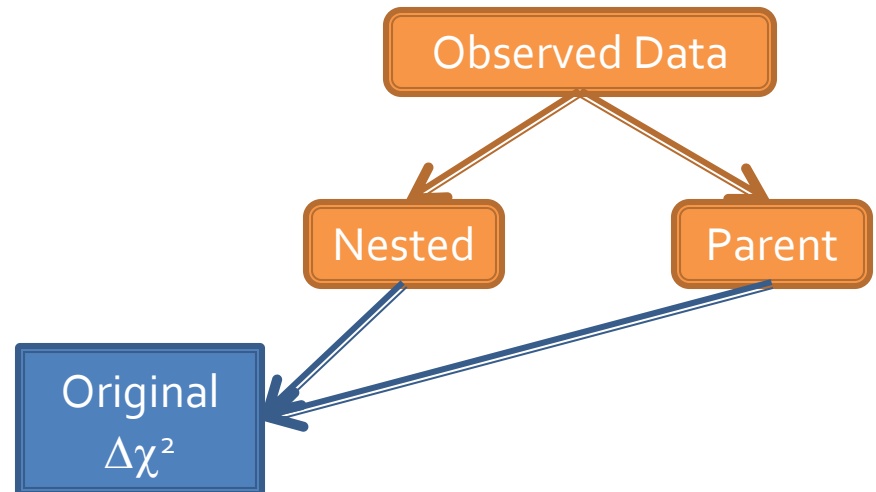
Parametric Bootstrap

- Use the parameter estimates from the nested model
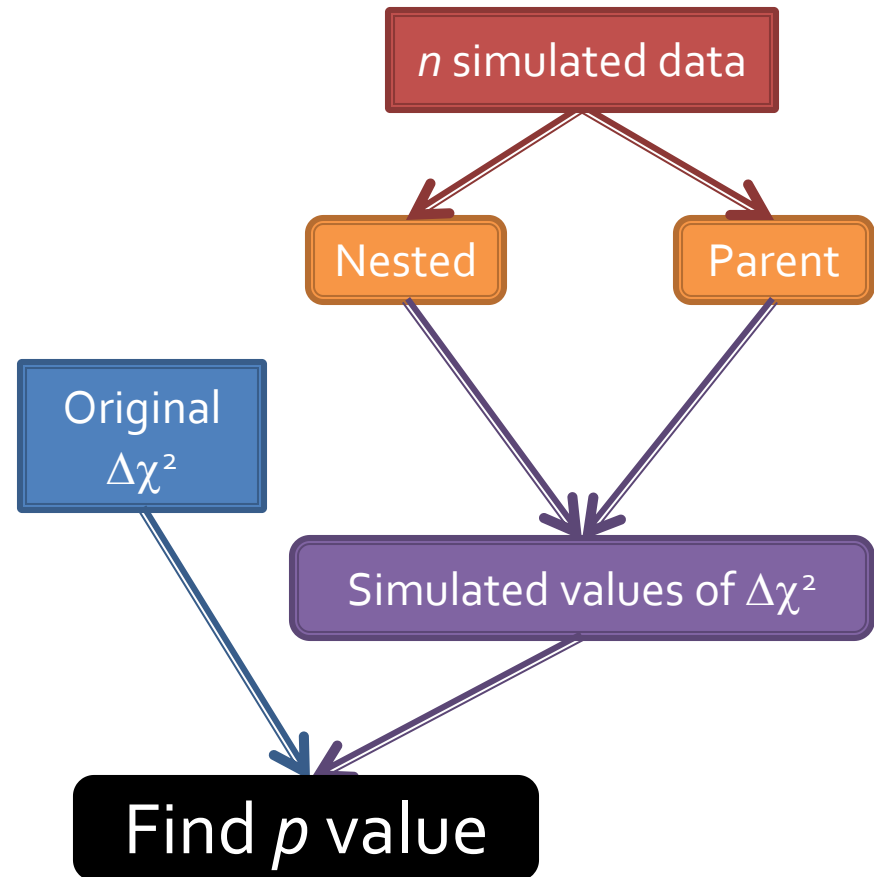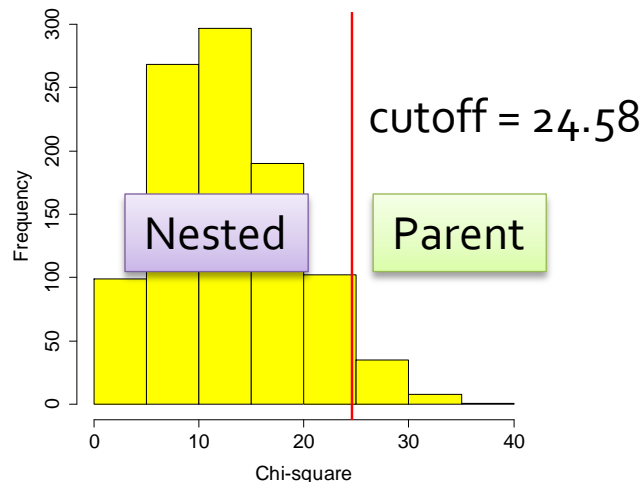- Add trivial misspecification
- Generate *n* datasets (e.g., 1,000)

Parametric Bootstrap

- Fit the nested and parent models to all simulated data
- Save the difference in a fit index for each simulated data
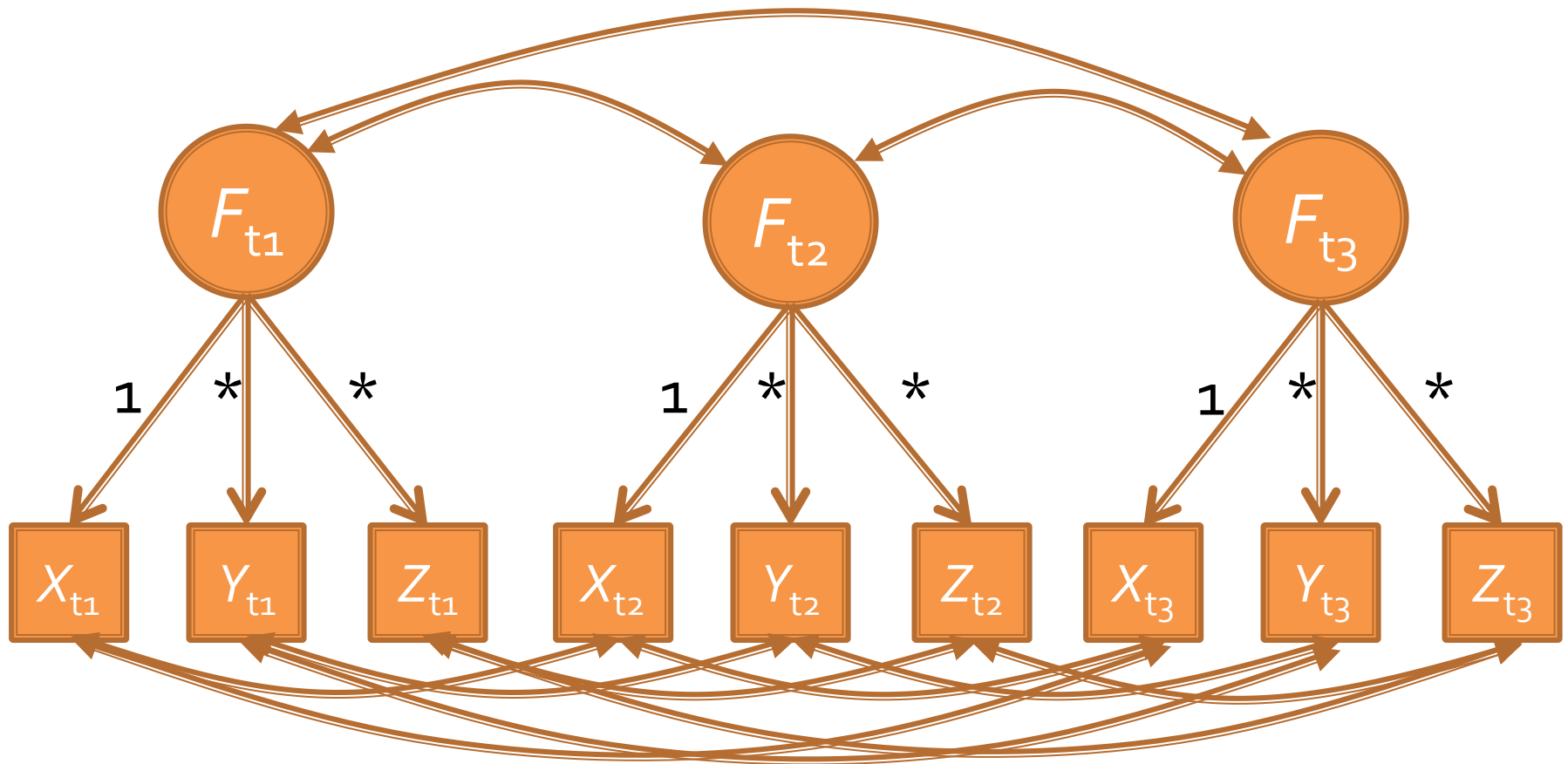- Plot the distribution of the difference

```
┌─────────────────────┐
│  n simulated data   │
└─────────────────────┘
     ↙           ↘
┌──────────┐  ┌──────────┐
│  Nested  │  │  Parent  │
└──────────┘  └──────────┘
     ↘           ↙
┌─────────────────────────────┐
│ Simulated values of Δχ²      │
└─────────────────────────────┘
```

Parametric Bootstrap

- Find *p* value
  - *p* ≤ α → select the parent model
  - *p* > α → select the nested model
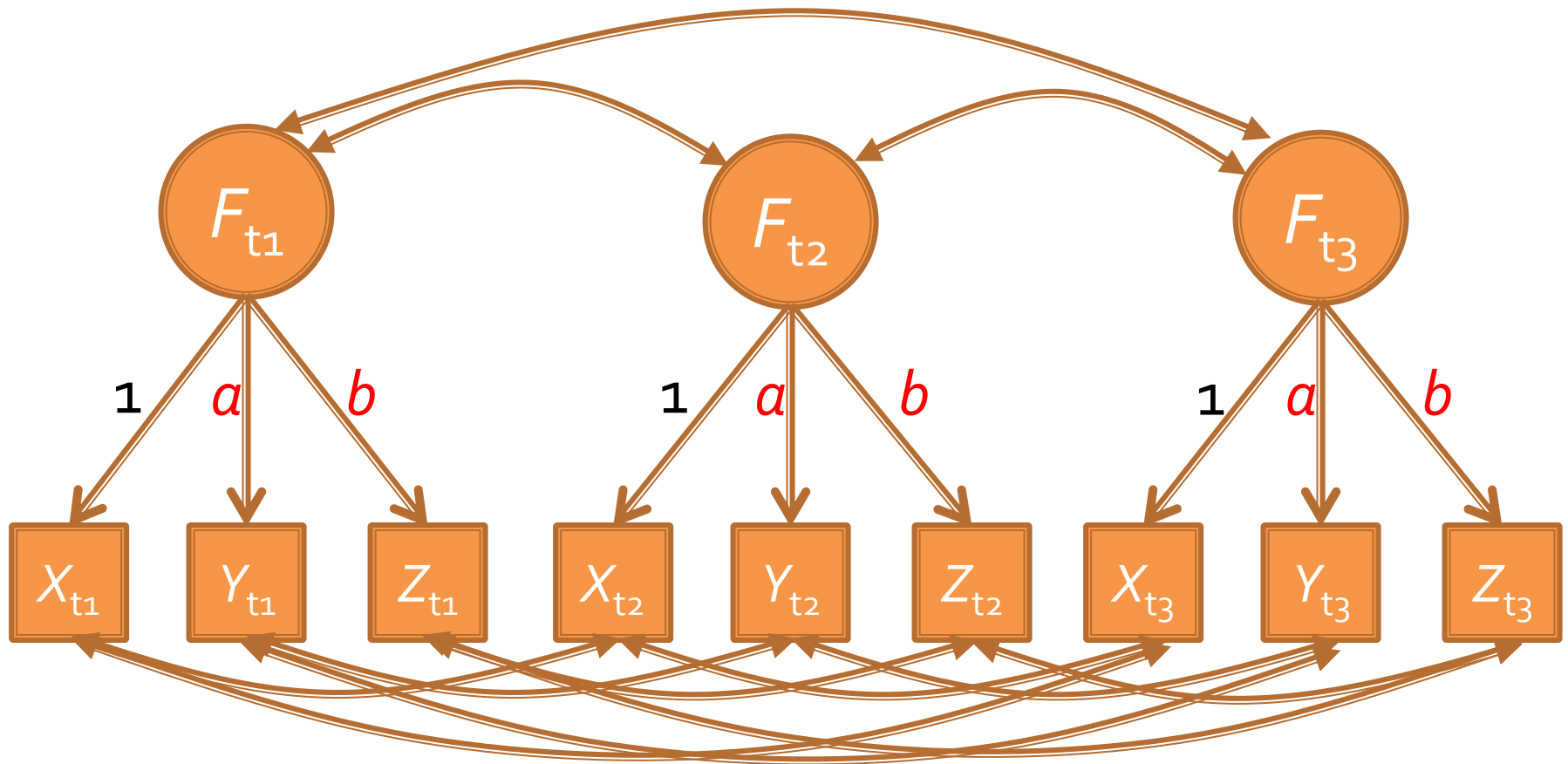

cutoff = 24.58

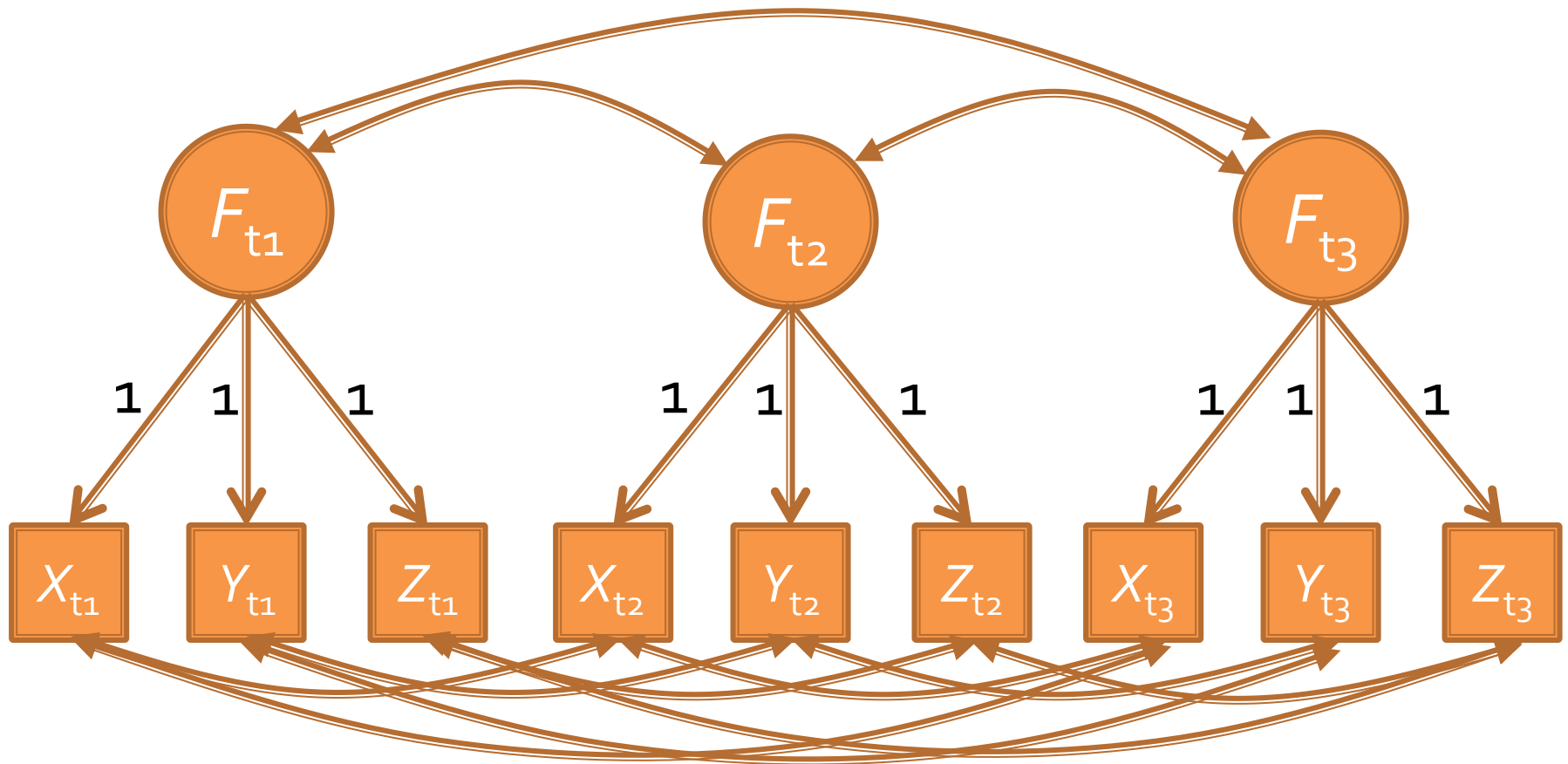# Simulation

Parent Model

Simulation

Nested Model

Simulation

- Sample size: 125, 250, 500, 1000
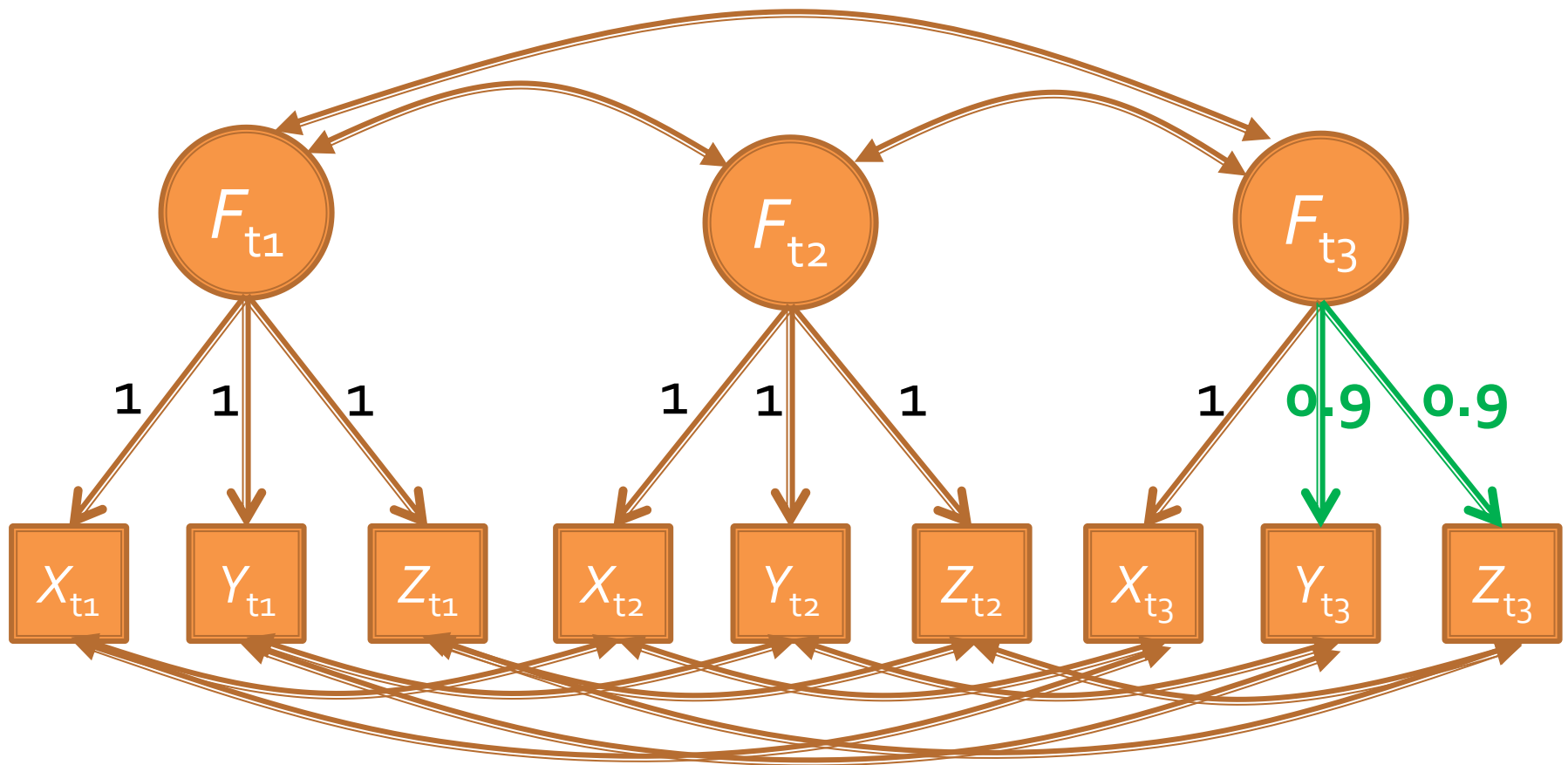- Severity of misspecification: None, Trivial, Severe

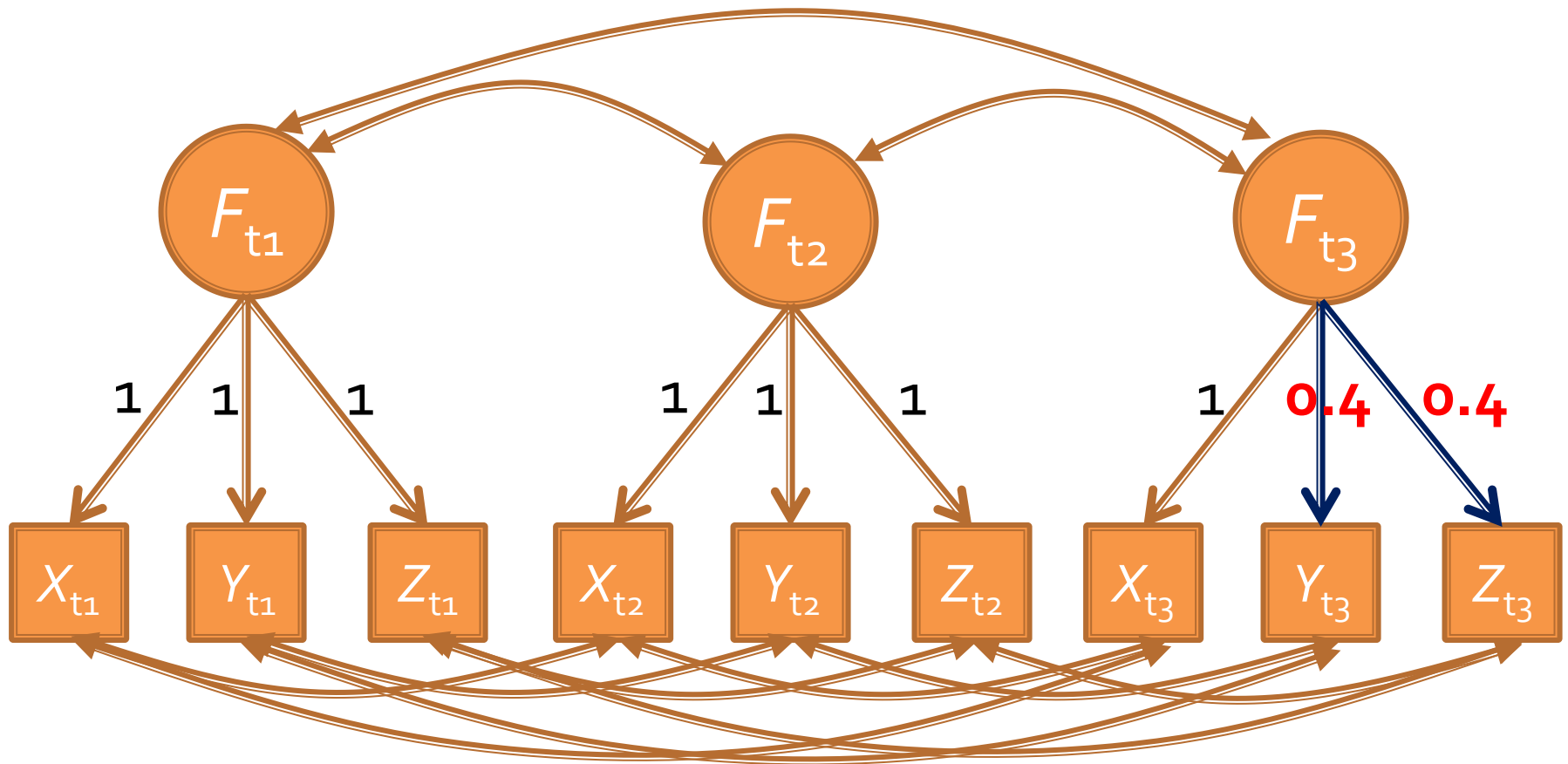# Simulation

Longitudinal Weak Invariance: No Misspecification

Simulation

Longitudinal Weak Invariance: Trivial Misspecification

Simulation

Longitudinal Weak Invariance: Severe Misspecification
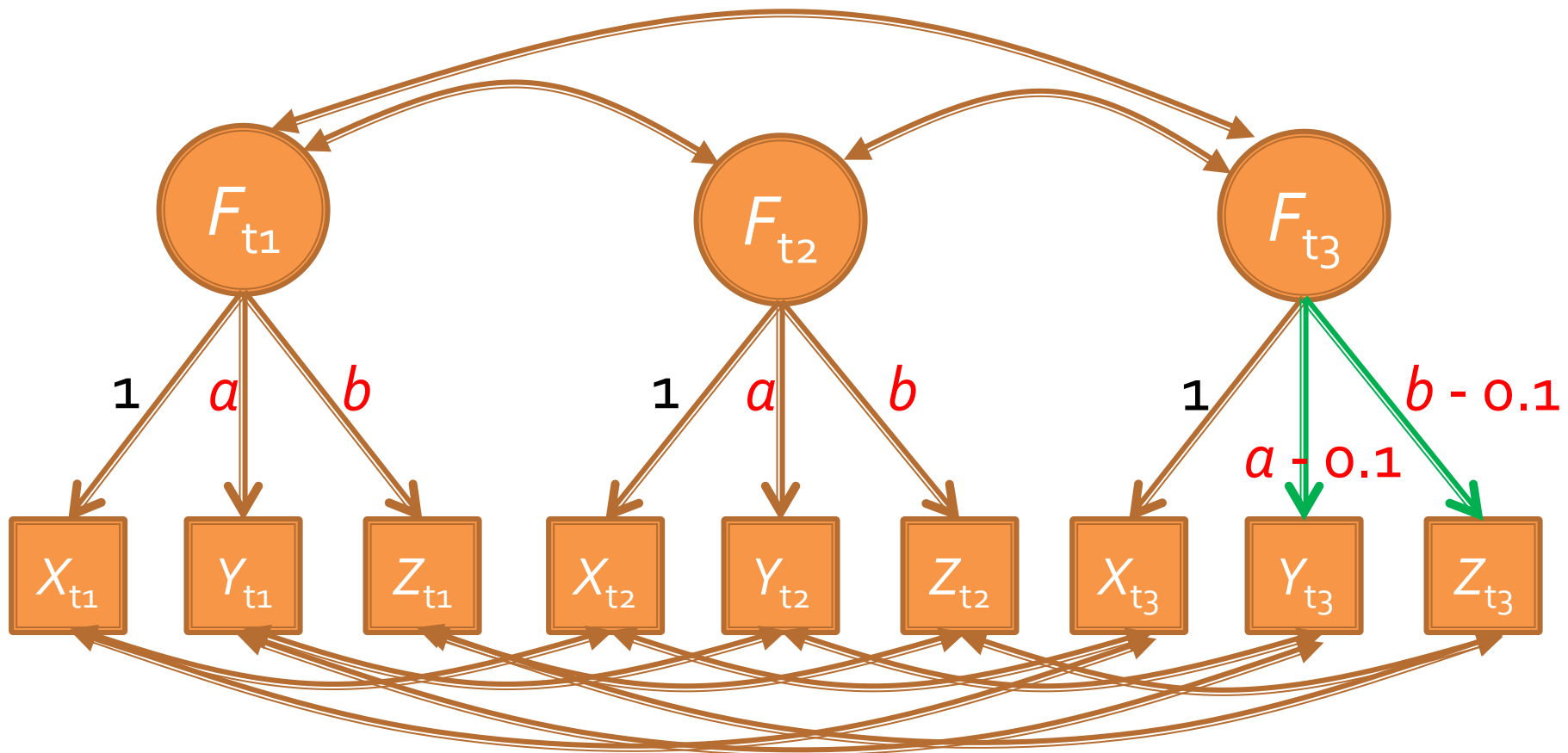
Simulation

- Sample size: 125, 250, 500, 1000
- Severity of misspecification: None, Trivial, Severe
- Three selection methods:
  - Chi-square difference test
  - Change in CFI < .002 (Meade et al., 2008)
  - Parametric Bootstrap
    - No misfit / Fixed / Random / Maximal
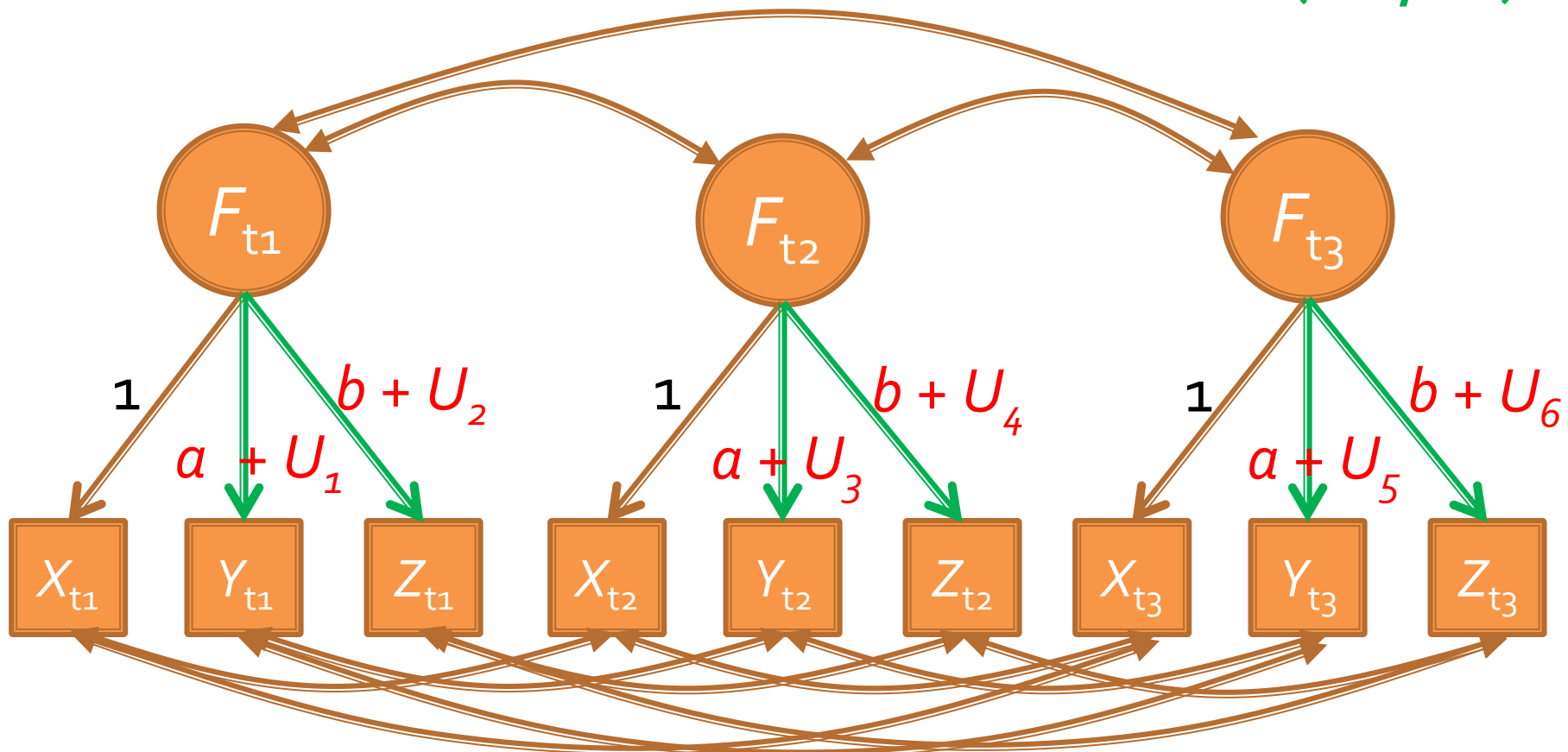
Simulation

Fixed Misspecification

Population RMSEA ≈ .045

Simulation

Random Misspecification

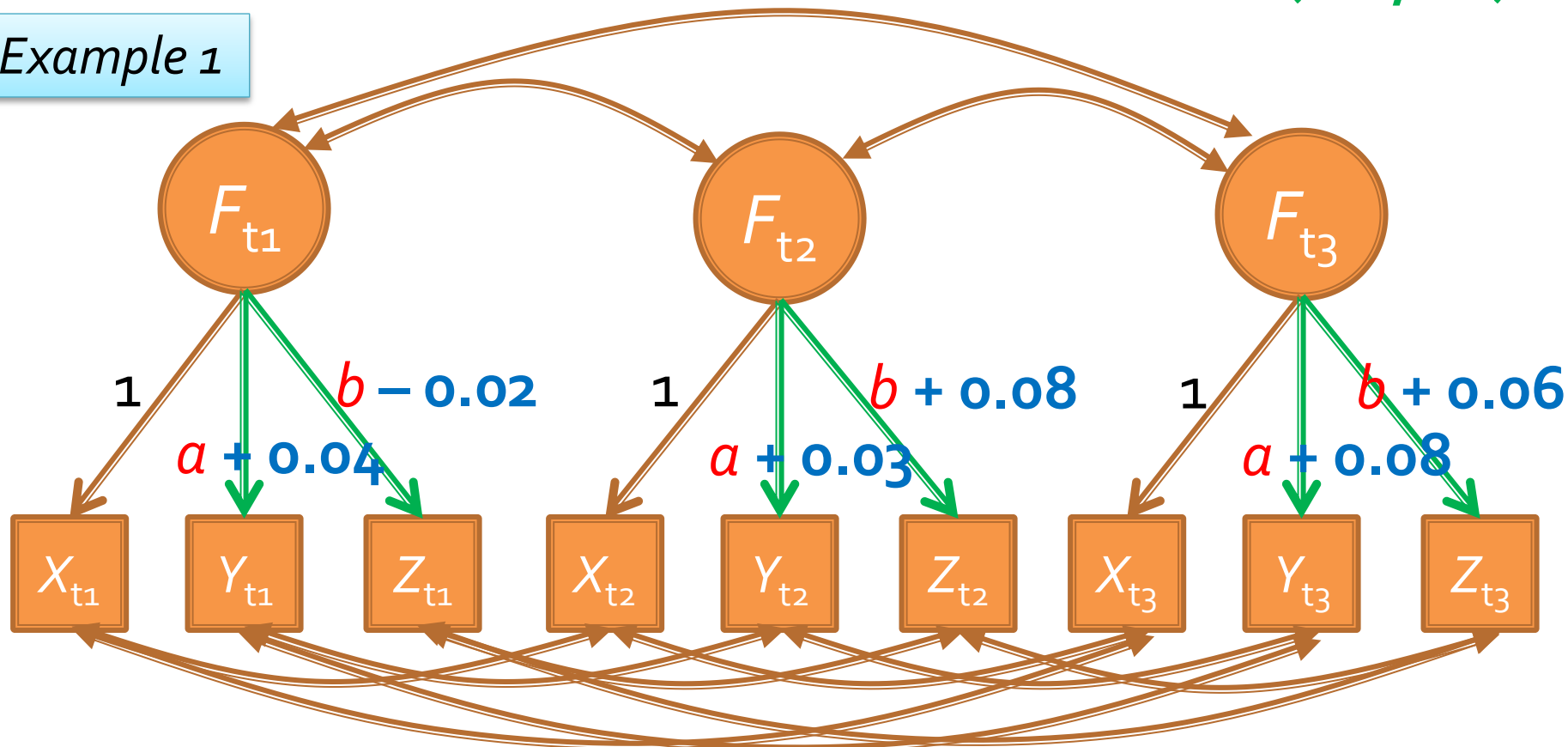U_1 to U_6 = A random draw from a uniform distribution (-0.1, 0.1)

Average Population RMSEA ≈ .051

Simulation

Random Misspecification

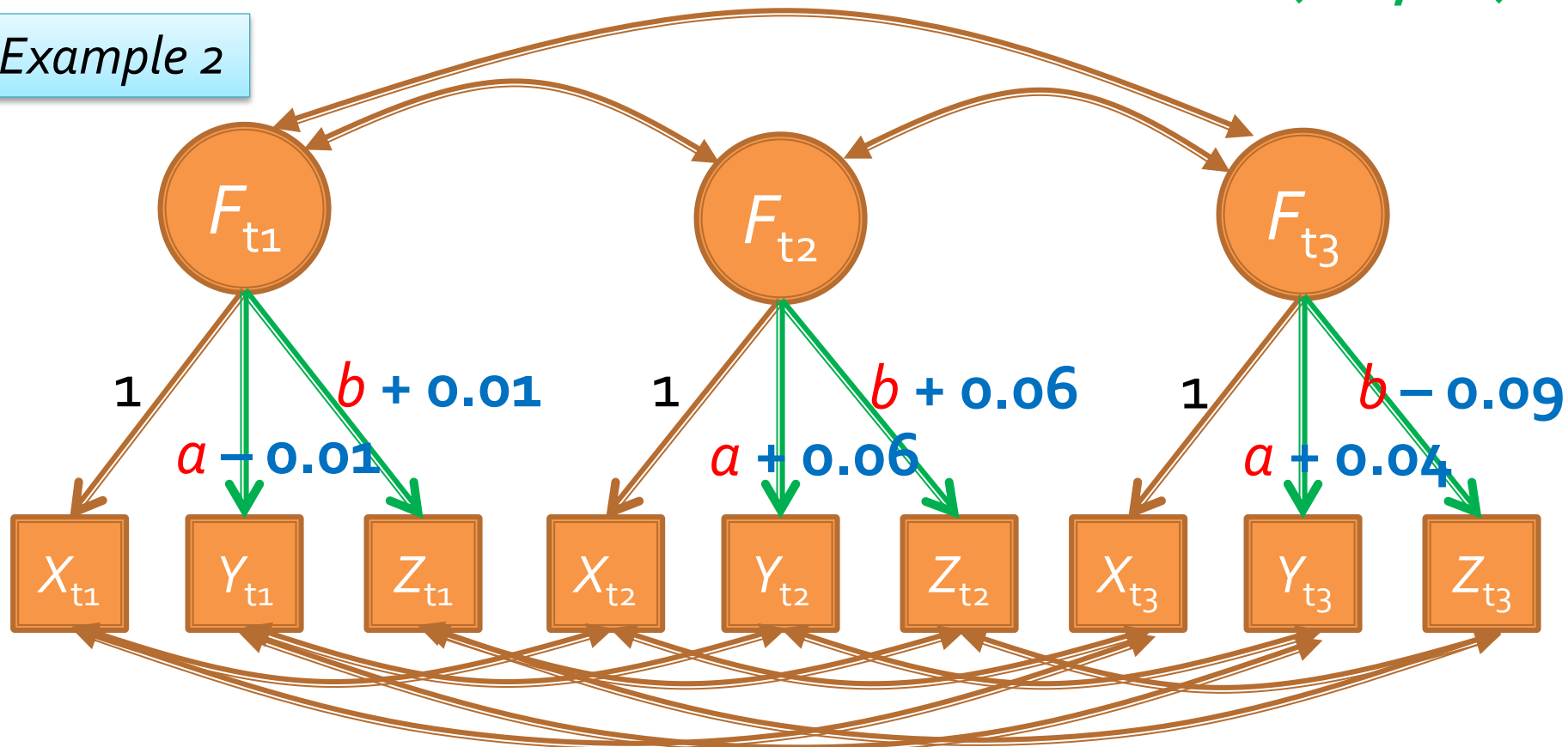U_1 to U_6 = A random draw from a uniform distribution (-0.1, 0.1)

Example 1

Average Population RMSEA ≈ .051

Simulation

Random Misspecification

U_1 to U_6 = A random draw from a uniform distribution (-0.1, 0.1)

Example 2

1    $b + 0.01$ 1 $b + 0.06$ 1 $b - 0.09$

$a - 0.01$ $a + 0.06$ $a + 0.04$

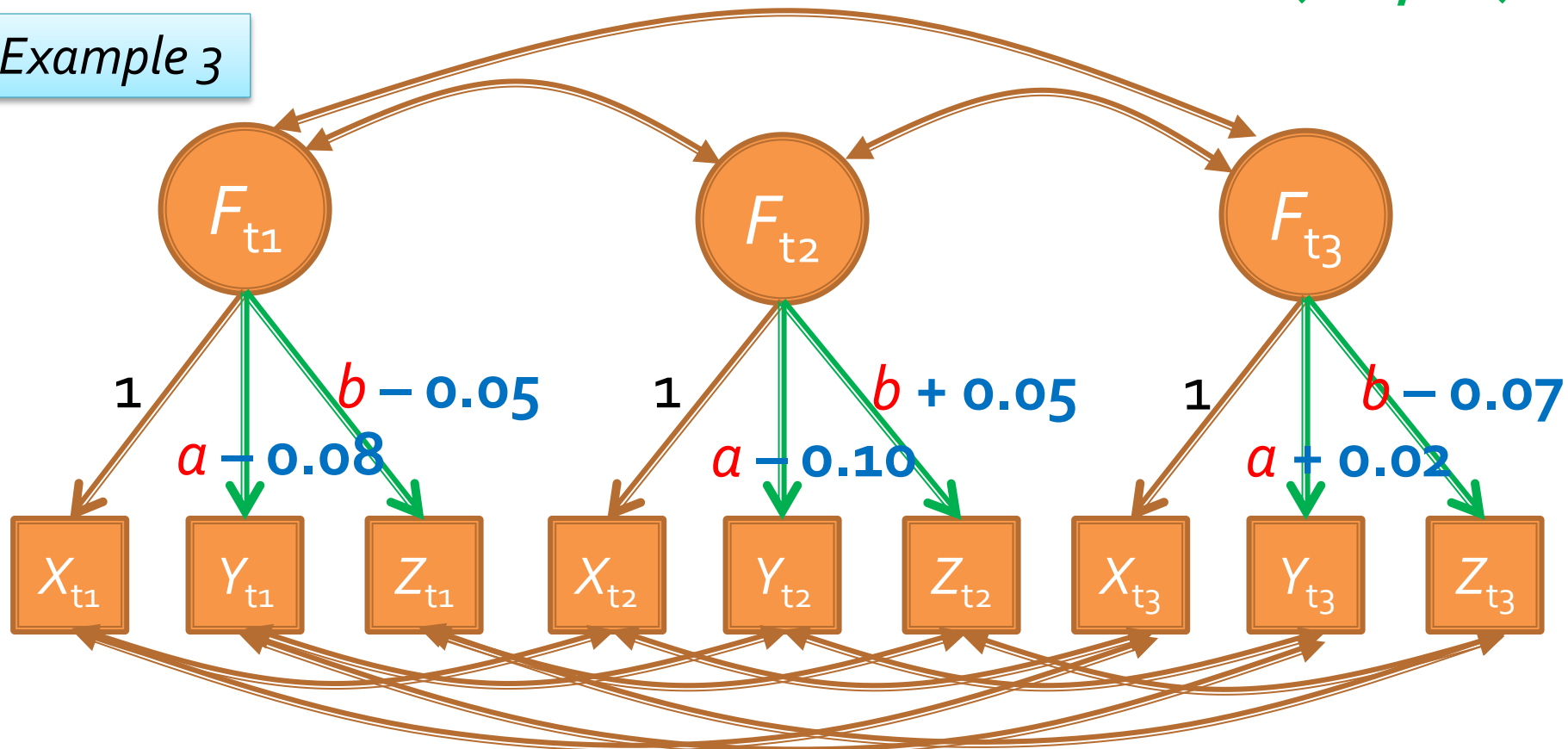Average Population RMSEA ≈ .051

Simulation

Random Misspecification

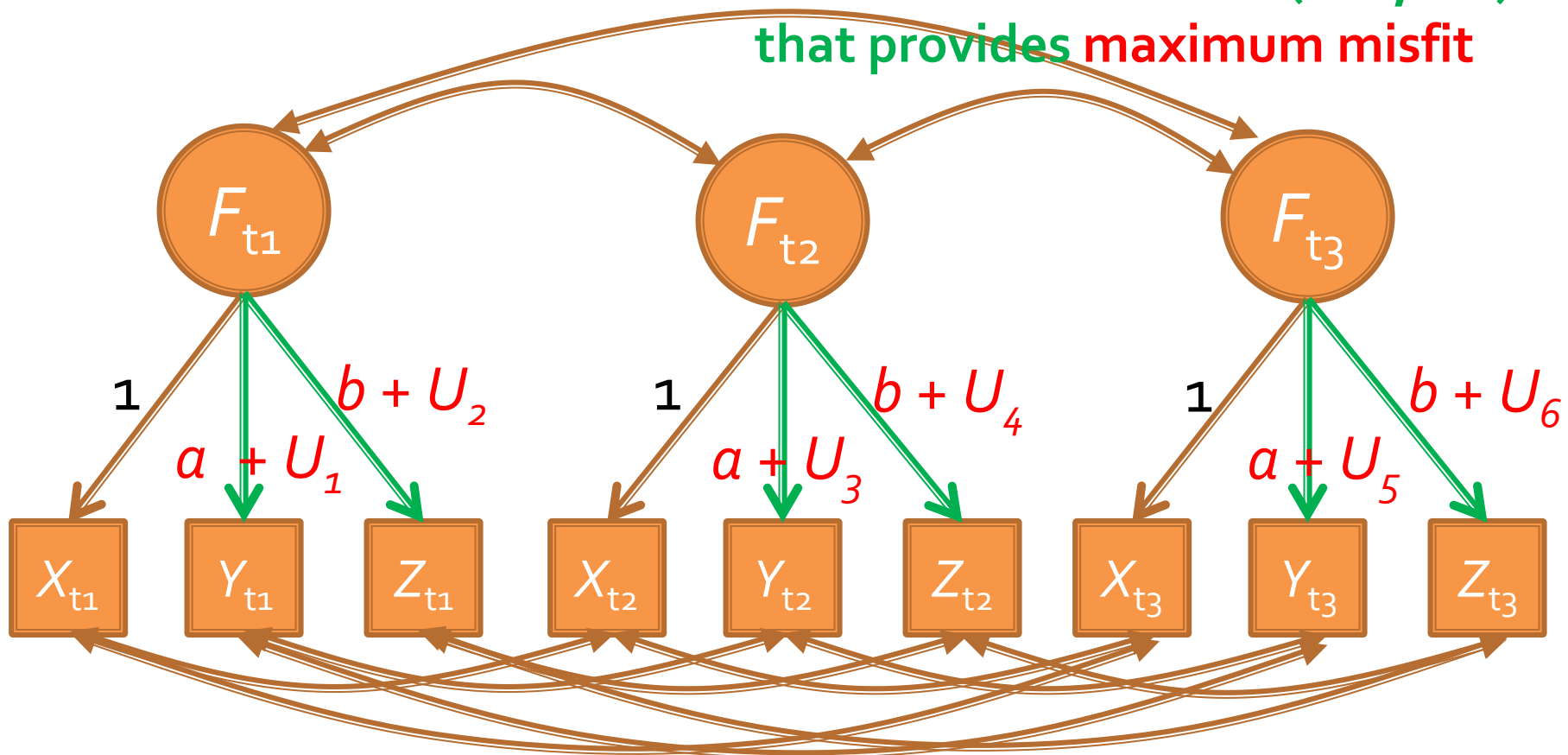U_1 to U_6 = A random draw from a uniform distribution (-0.1, 0.1)

Example 3

F_{t1} F_{t2} F_{t3}

1    $b - 0.05$ 1 $b + 0.05$ 1 $b - 0.07$

$a - 0.08$ $a - 0.10$ $a + 0.02$

X_{t1} Y_{t1} Z_{t1} X_{t2} Y_{t2} Z_{t2} X_{t3} Y_{t3} Z_{t3}

Average Population RMSEA ≈ .051

Simulation

Maximal Misspecification

U_1 to U_6 = An only **set of draw** from **a uniform distribution (-0.1, 0.1)** that provides **maximum misfit**

Population RMSEA ≈ .08

Simulation

Maximal Misspecification

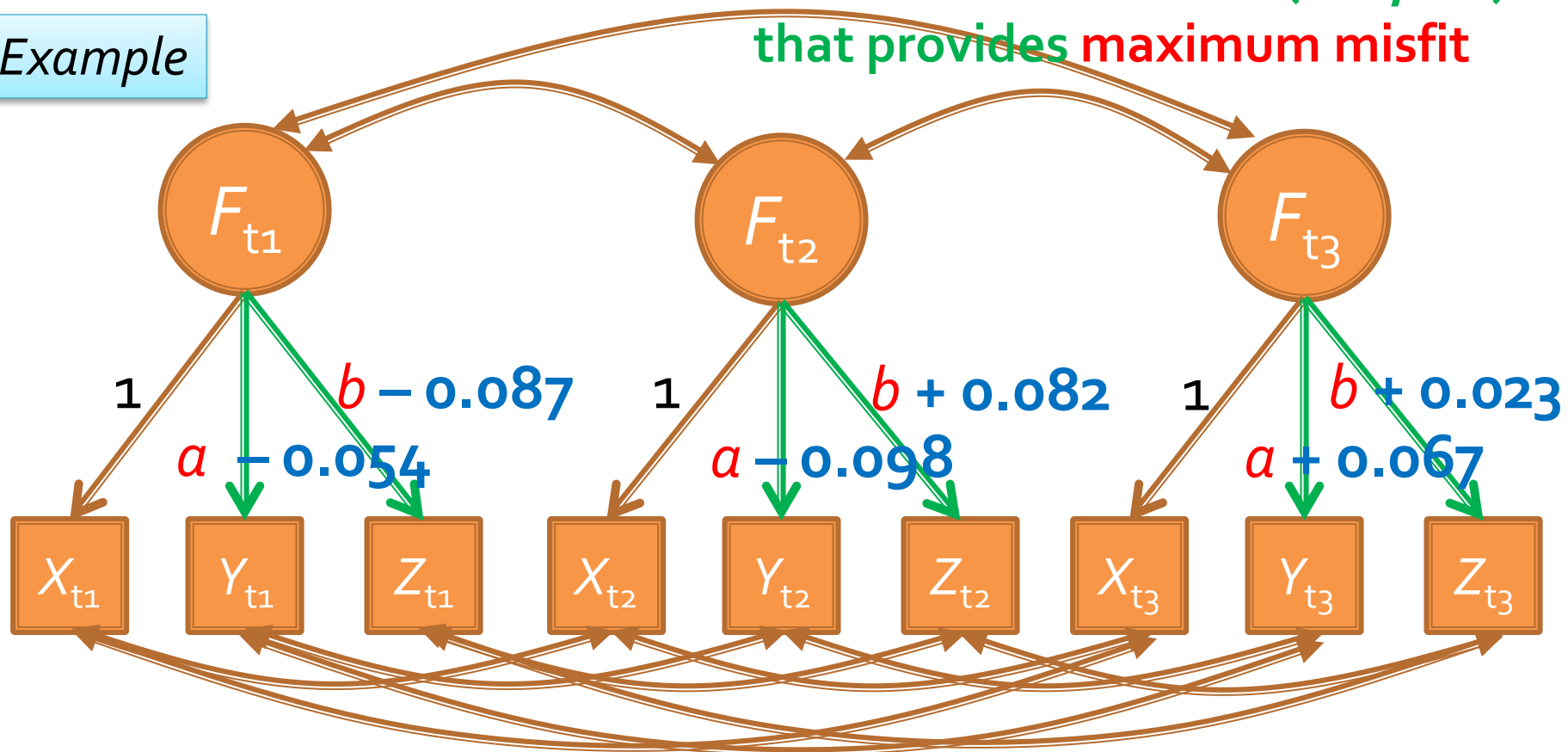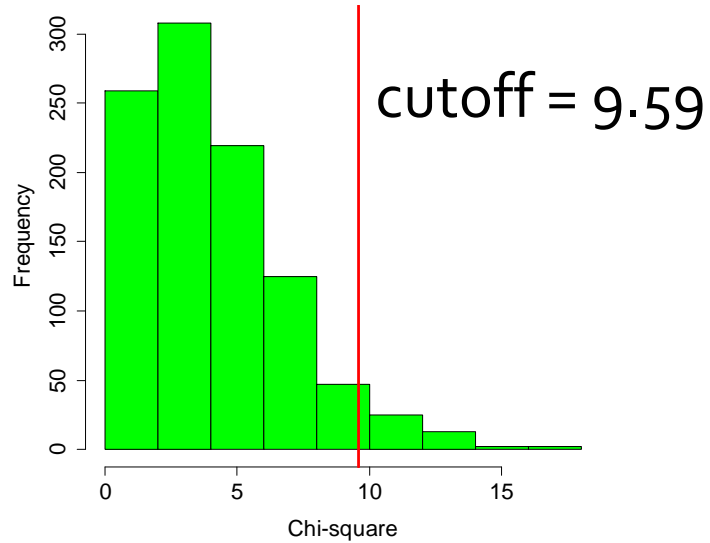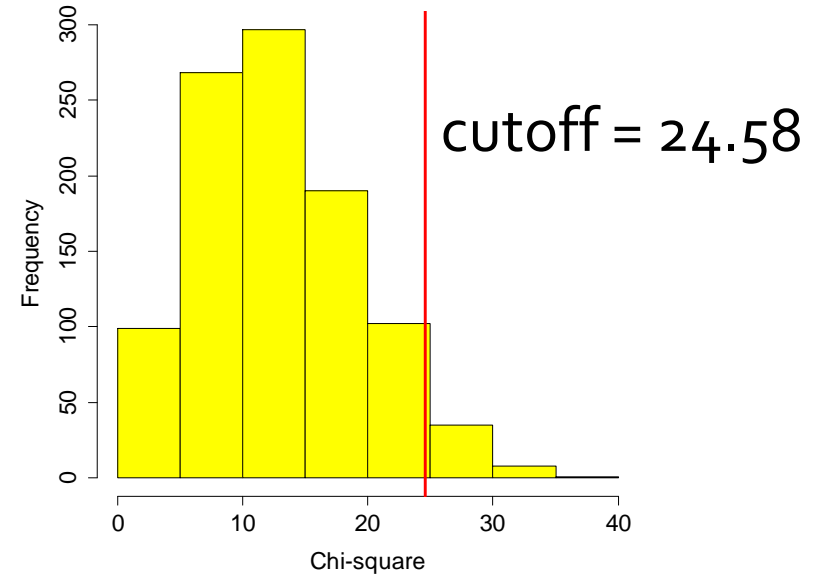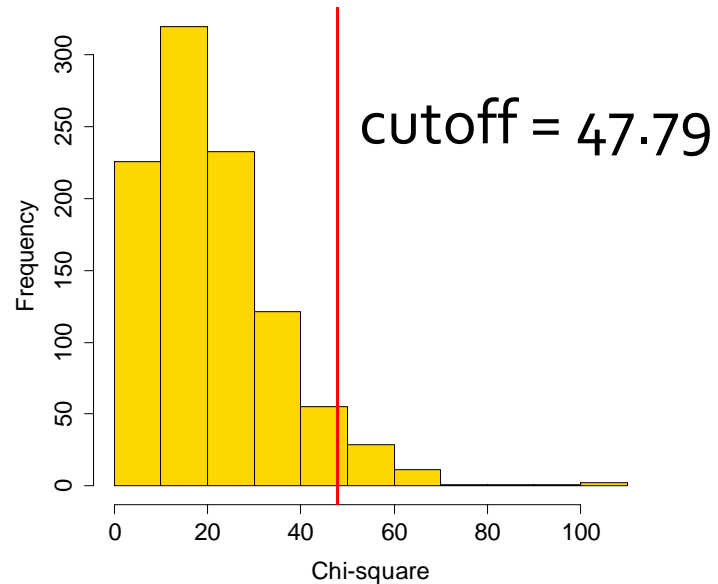U_1 to U_6 = An only set of draw from a uniform distribution (-0.1, 0.1) that provides maximum misfit

Example

$b - 0.087$ $b + 0.082$ $b + 0.023$

$a - 0.054$ $a - 0.098$ $a + 0.067$

Population RMSEA ≈ .08

No Misspecification — cutoff = 9.59

Fixed — cutoff = 24.58
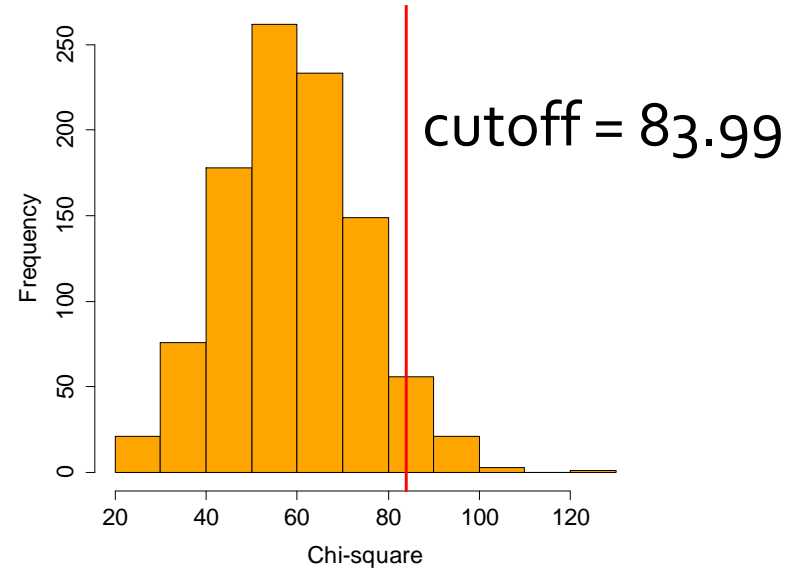
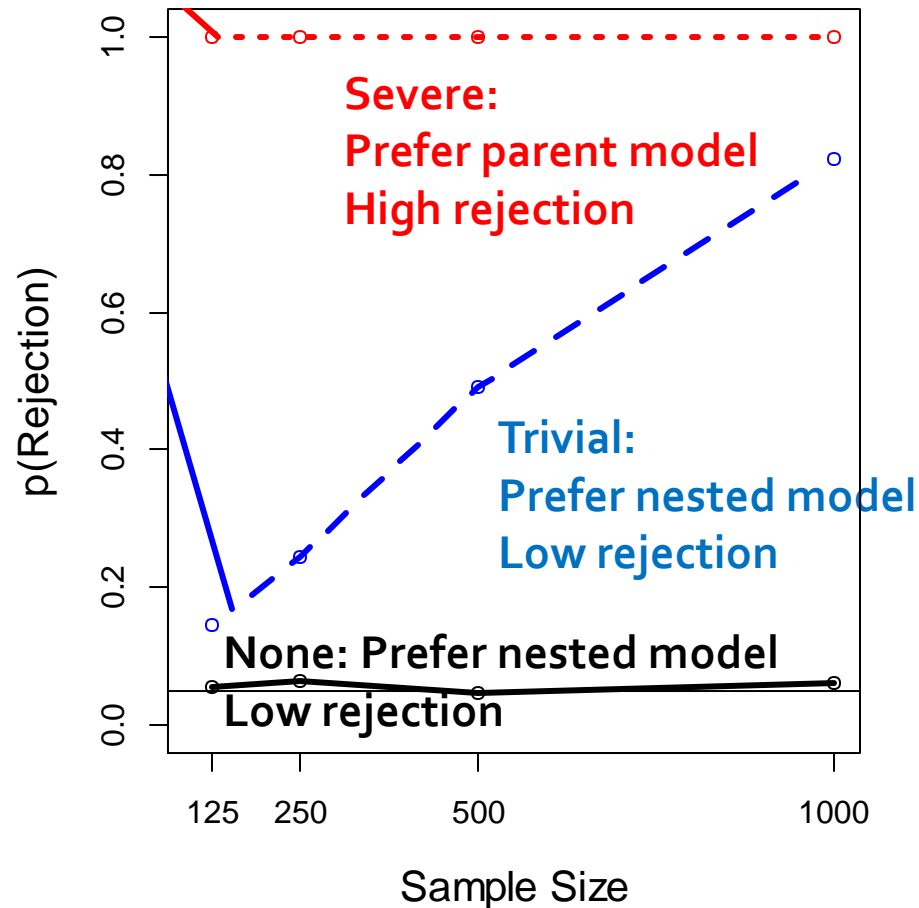**Random** — cutoff = 47.79

Maximal — cutoff = 83.99
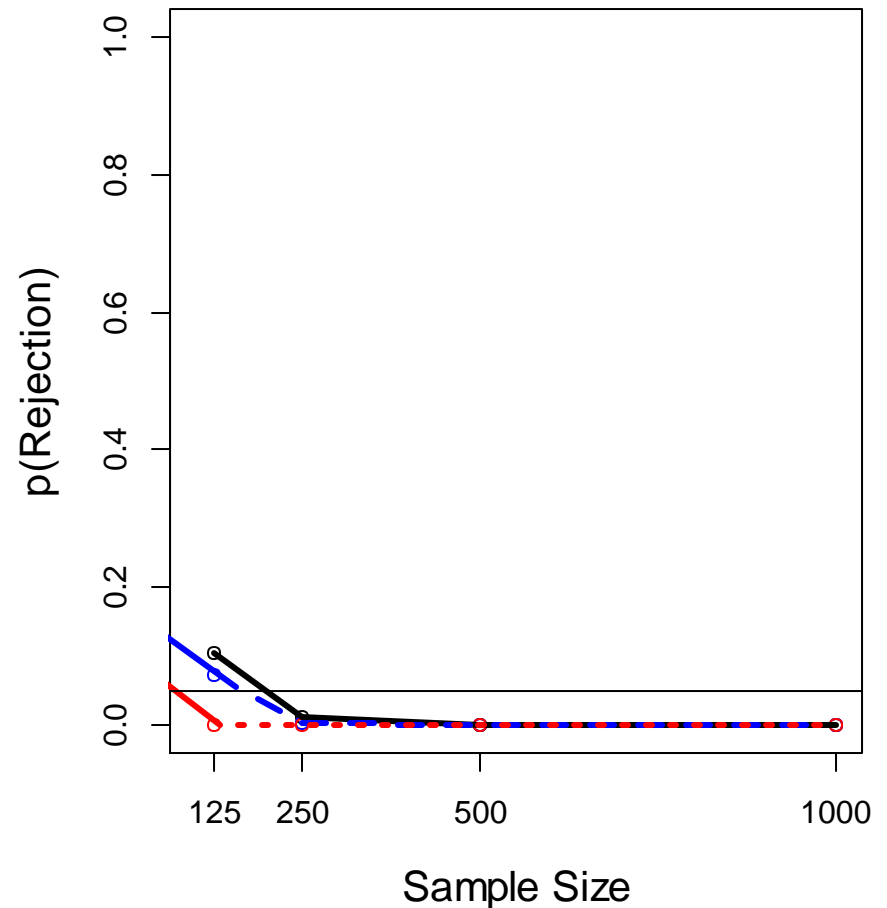
Simulation

- 1,000 Replications
- Rejection Rate
  - Low when population misfit is none or trivial
  - High when population misfit is severe
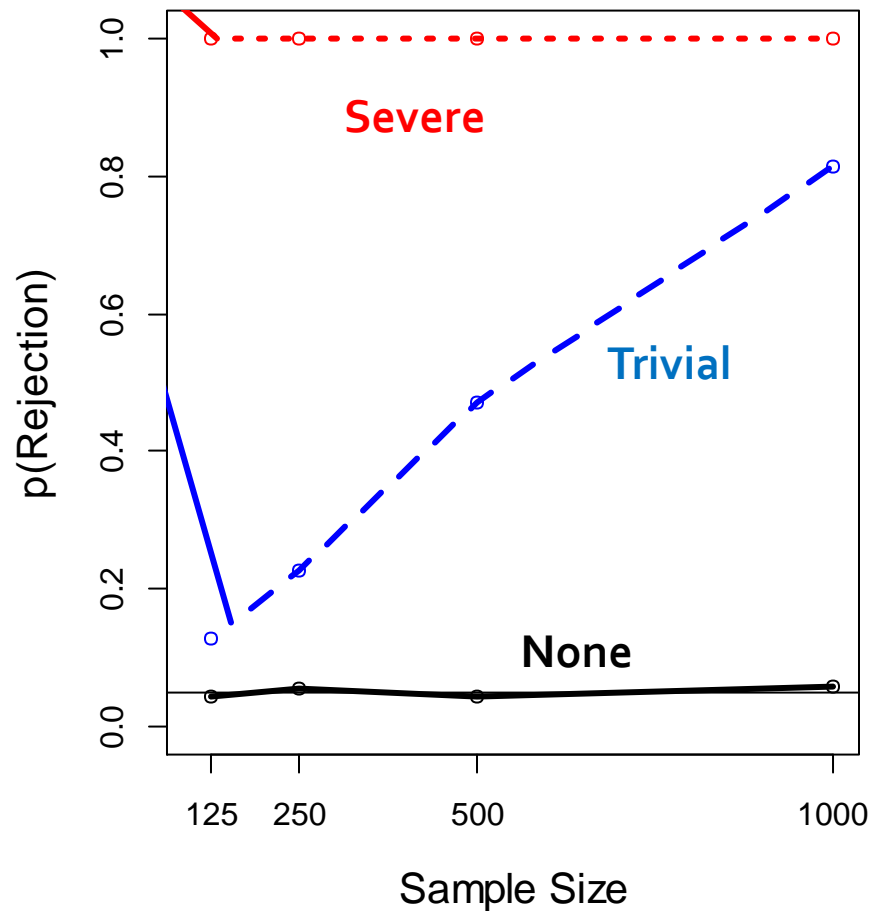
Result

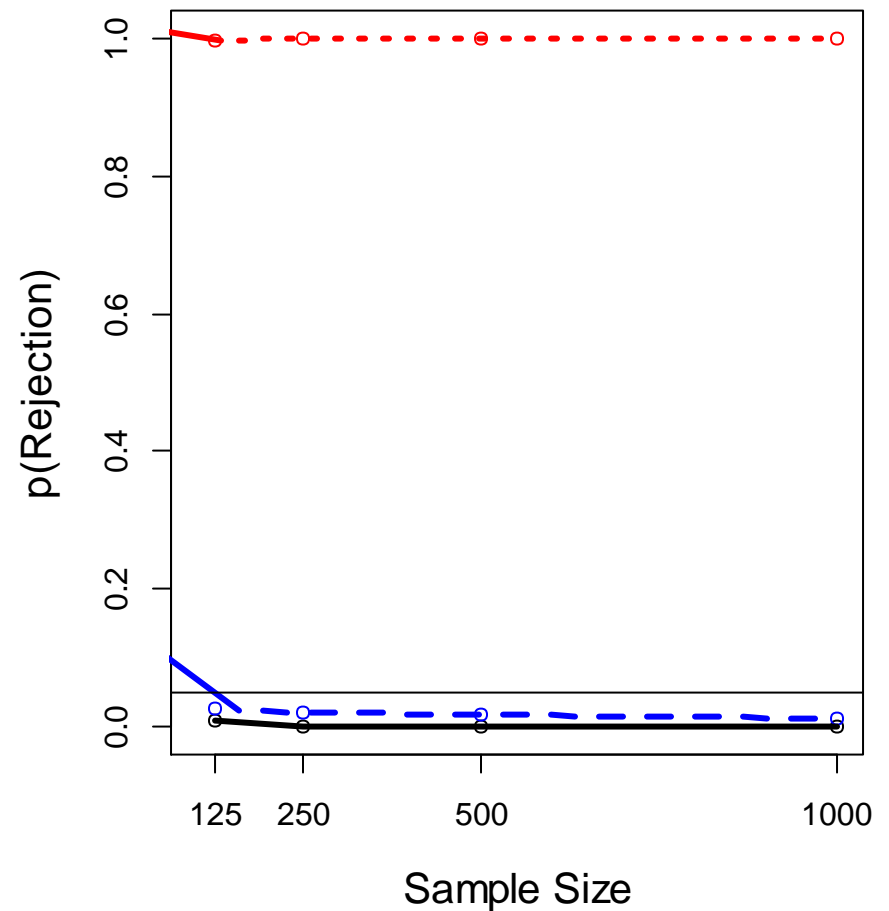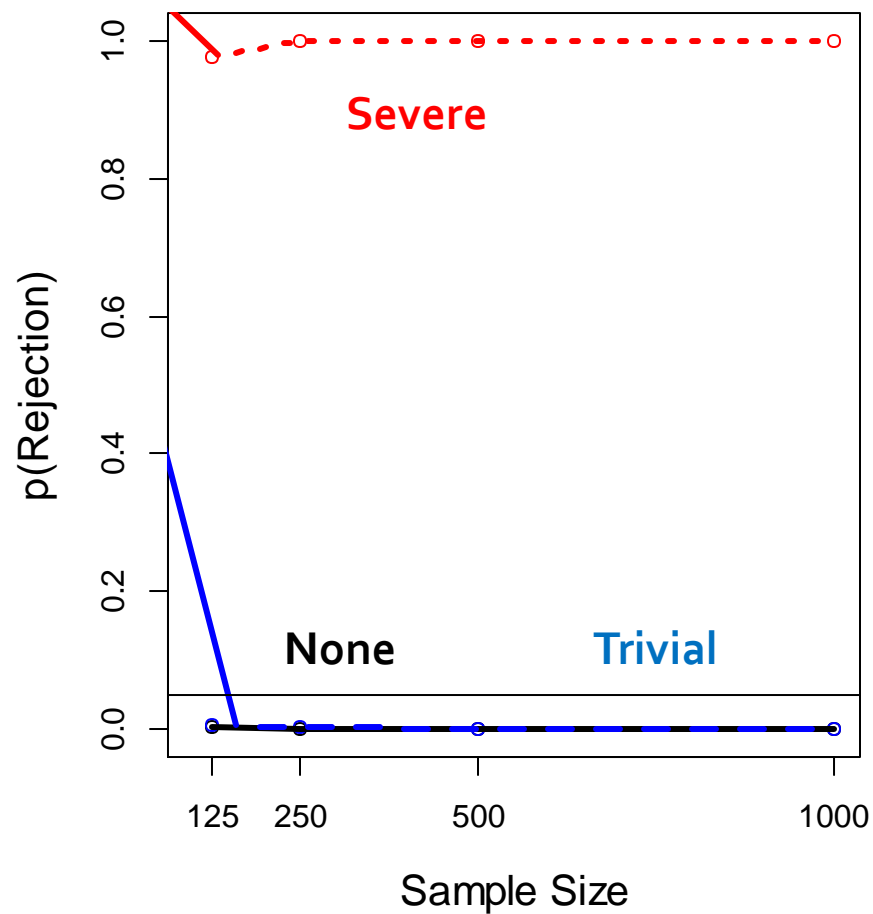**Chi-square difference test**

Change in CFI

Severe:
Prefer parent model
High rejection

Trivial:
Prefer nested model
Low rejection

None: Prefer nested model
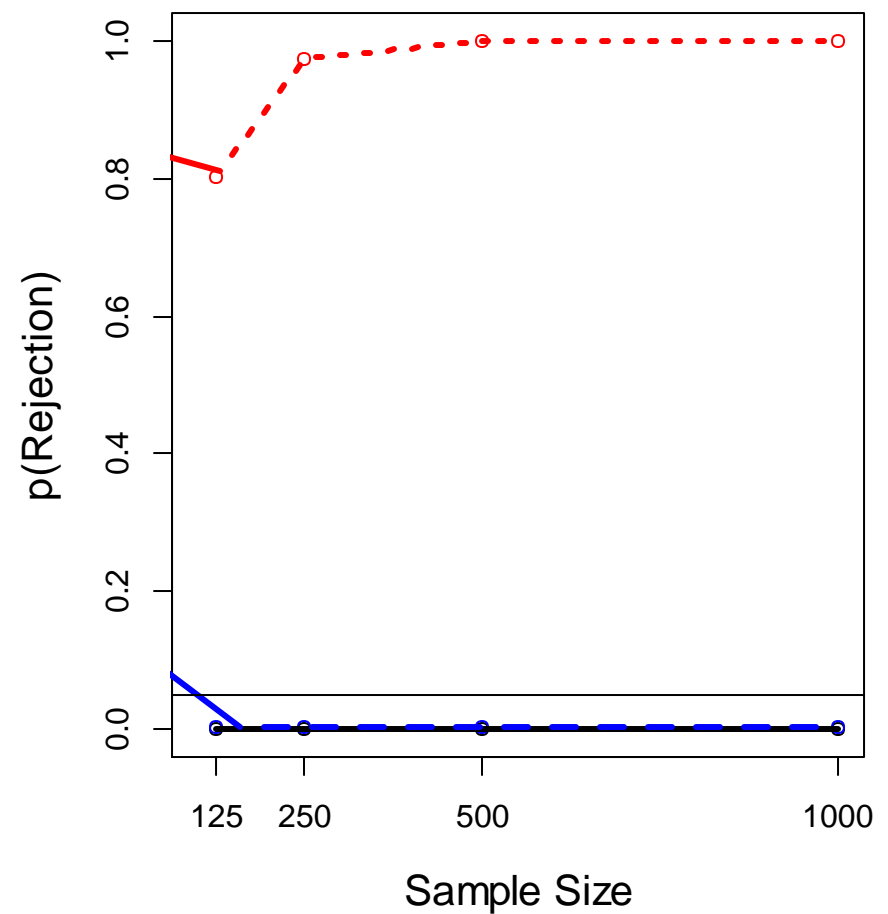Low rejection

Result

Result

PB: Random method with uniform misfit

PB: Maximal method

Discussion

- Chi-square test = Parametric bootstrap without trivial misspecification, asymptotically
  - Reject trivial misfit in large sample size
- CFI cutoff of .002 (or .01) has extremely low power
- Parametric bootstrap with defined trivial misspecification is good
  - Avoid universal cutoff criterion
  - Avoid rejection of trivial misspecification
  - Encourage Users to define trivial misspecification

Discussion

- Can develop your own simulation in R
- Parametric bootstrap approach in absolute fit and nonnested model comparison
- Semi-parametric bootstrap
- This procedure can be implemented in the *simsem* package in R
https://github.com/simsem/simsem/wiki

Acknowledgement

KU CENTER FOR RESEARCH METHODS & DATA ANALYSIS
College of Liberal Arts & Sciences

Thank you!!
Questions, Comments, Concerns