# วิธีการแบบบูรณาการในการประเมินความเหมาะสมของโมเดลใน Structural Equation Modeling

สันทัด พรประเสริฐมานิต

Wei Wu

# โครงร่างการนำเสนอ

- วิธีการที่ใช้ในปัจจุบัน
- ปัญหาของวิธีการที่ใช้ในปัจจุบัน
- วิธีการอื่นที่พัฒนาขึ้น
- วิธีการแบบบูรณาการ
- การจำลองสถานการณ์
- สรุป

# วิธีการที่ใช้ในปัจจุบัน

- ใน Structural Equation Modeling ผู้วิเคราะห์จำเป็นต้องเลือกที่จะใช้หรือปฏิเสธโมเดล
- ความเหมาะสมในภาพรวม (Global fit): ใช้จุดตัดของดัชนีความเหมาะสม (Fit indices)
  - RMSEA: <= .06
  - CFI: >= .95
  - TLI: >= .95
  - SRMR: <= .08
- ความเหมาะสมจำเพาะ (Local fit): ตรวจสอบโดยดัชนีการปรับโมเดล (Modification Indices)

# ดัชนีความเหมาะสมเป็นการวัดขนาดอิทธิพล

- เป็นวิธีวัดขนาดอิทธิพลรูปแบบหนึ่ง กล่าวคือ วัดขนาดความแตกต่างระหว่าง
  - Observed mean vector and covariance matrix
  - Model-implied mean vector and covariance matrix
- คุณสมบัติที่ดีของขนาดอิทธิพล คือ ค่าไม่เปลี่ยนแปลงหากจำนวนกลุ่มตัวอย่างเปลี่ยนแปลง
- ค่าพื้นฐานที่ใช้ในการคำนวณดัชนีทั้งหมด คือ Discrepancy value ($F_M$)

$$F_M = \mathrm{tr}(\Sigma\Sigma_M^{-1}) - \log\left|\Sigma\Sigma_M^{-1}\right| + (\mu - \mu_M)'\Sigma_M^{-1}(\mu - \mu_M)$$

# ดัชนีความเหมาะสมเป็นการวัดขนาดอิทธิพล

- ดัชนีความเหมาะสมแบบสัมบูรณ์ (Absolute fit indices)
  - RMSEA

$$\varepsilon_M = \sqrt{\frac{F_M}{df_M}}$$

  - SRMR

$$\text{SRMR} = \sqrt{\frac{\sum_j \sum_{k \le j} \left(\rho_{Mjk} - \rho_{jk}\right)^2}{\frac{p(p+1)}{2}}}$$

# ดัชนีความเหมาะสมเป็นการวัดขนาดอิทธิพล

- ดัชนีความเหมาะสมแบบสัมบูรณ์ (Absolute fit indices)
  - CFI

$$\mathrm{CFI} = 1 - \frac{F_M}{F_0}$$

  - TLI (หรือ NNFI)

$$\mathrm{TLI} = 1 - \frac{F_M}{F_0} \cdot \frac{df_0}{df_M}$$

# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

- ขนาดของโมเดล
- ชนิดของโมเดล
- ขนาดของค่า Parameter ในโมเดล

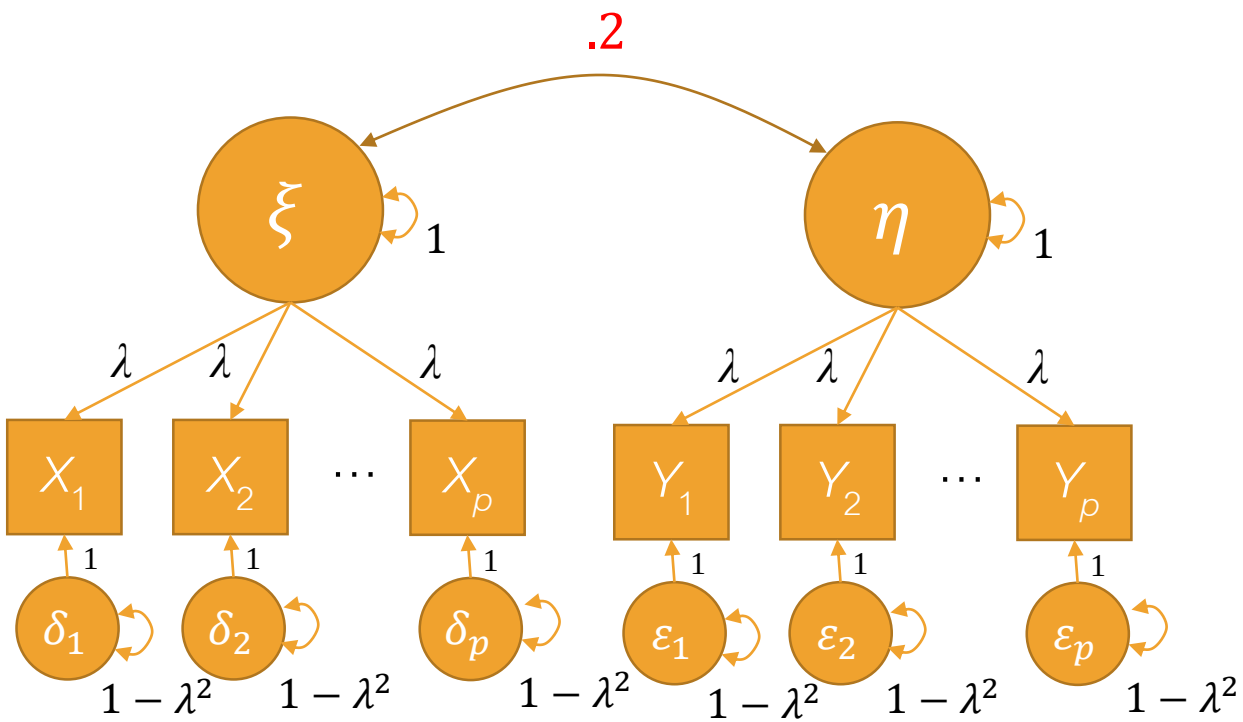
คุณสมบัติของโมเดล

- ขนาดกลุ่มตัวอย่างและข้อมูลสูญหาย
- วิธีการประมาณค่า
- การกระจายของข้อมูล

คุณสมบัติของการข้อมูล
และการวิเคราะห์ข้อมูล

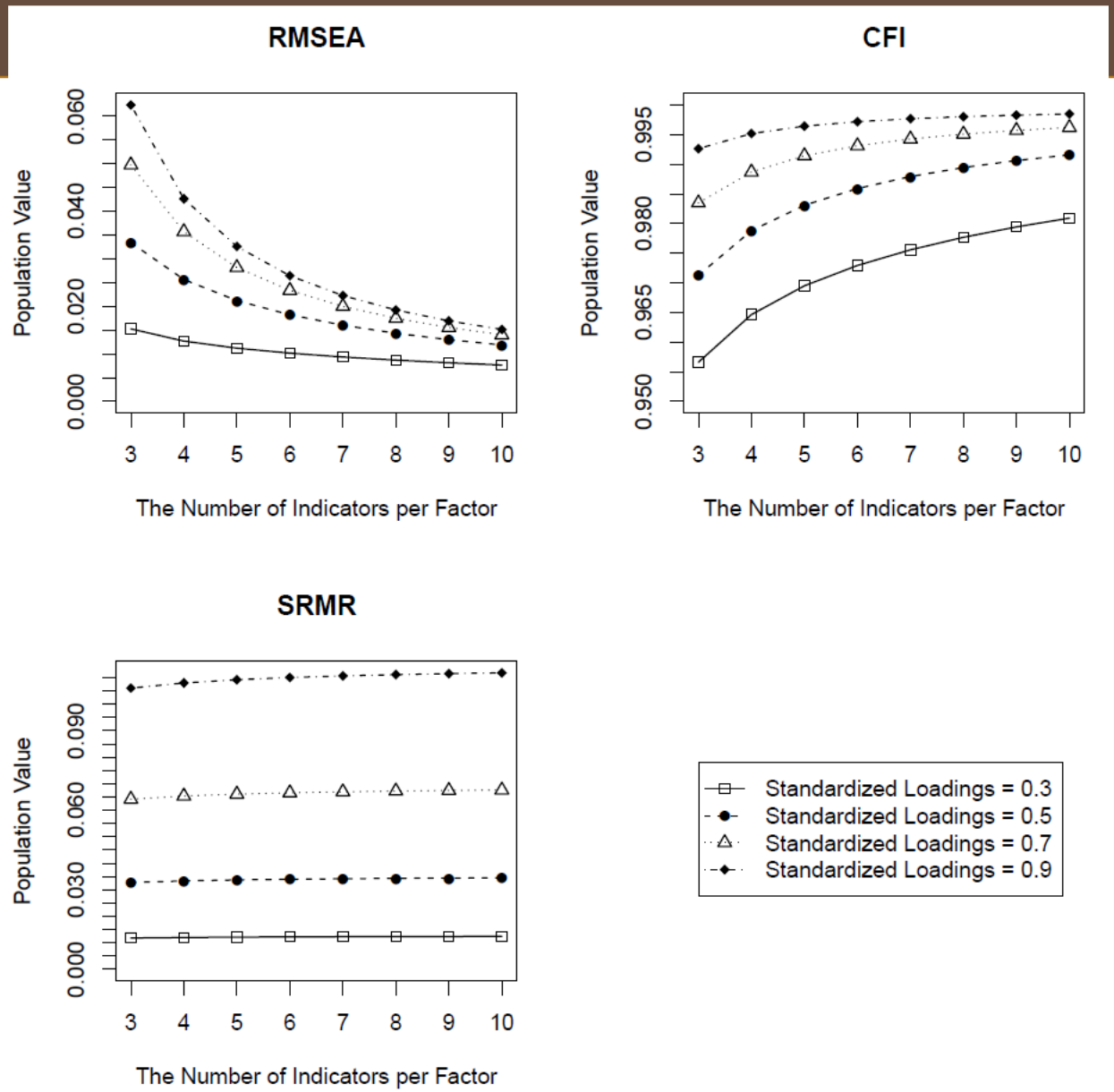
# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

- ขนาดของโมเดล
- ขนาดของค่า Parameter ในโมเดล



# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

- ขนาดของโมเดล
- ขนาดของค่า Parameter ในโมเดล



# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

- **ชนิดของโมเดล**
  - งานวิจัยในอดีต หาจุดตัดของดัชนีเหล่านี้จาก Confirmatory Factor Analysis ซึ่งเกี่ยวข้องกับการอธิบายความแปรปรวนร่วมเท่านั้น
  - ในปัจจุบันพบว่าจุดตัดที่ได้ ไม่สามารถใช้ได้กับโมเดลอื่น เช่น Growth Curve Analysis (Wu, 2008) หรือ Multilevel Confirmatory Factor Analysis (Boulton, 2011)

# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

- **จำนวนกลุ่มตัวอย่างและข้อมูลสูญหาย**
  - ดัชนีที่ได้รับความนิยม เช่น RMSEA, CFI, TLI ไม่ได้รับผลกระทบจากจำนวนกลุ่มตัวอย่าง ยกเว้นแต่จำนวนกลุ่มตัวอย่างจะน้อยมาก (< 200)
  - ดัชนีที่ไม่ได้รับความนิยม เช่น GFI, AGFI, RMR ได้รับผลกระทบจากจำนวนกลุ่มตัวอย่าง
  - จำนวนข้อมูลสูญหายมากขึ้น ดัชนีเหล่านี้จะบอกว่าโมเดลเหมาะสมกว่าความเป็นจริง (Davey, 2005)
  - ความสัมพันธ์นี้จะมากขึ้น หากรูปแบบข้อมูลสูญหายเป็นแบบ Missing at random (Davey, 2005)

# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

- **วิธีการประมาณค่า**
  - ค่า $\chi^2$ หรือ $F_M$ จะมีค่าแตกต่างกัน หากใช้วิธีการประมาณค่าที่แตกต่างกัน (ML, ADF, GLS)

# ปัจจัยที่ส่งผลให้ดัชนีของเหมาะสมเปลี่ยนแปลง

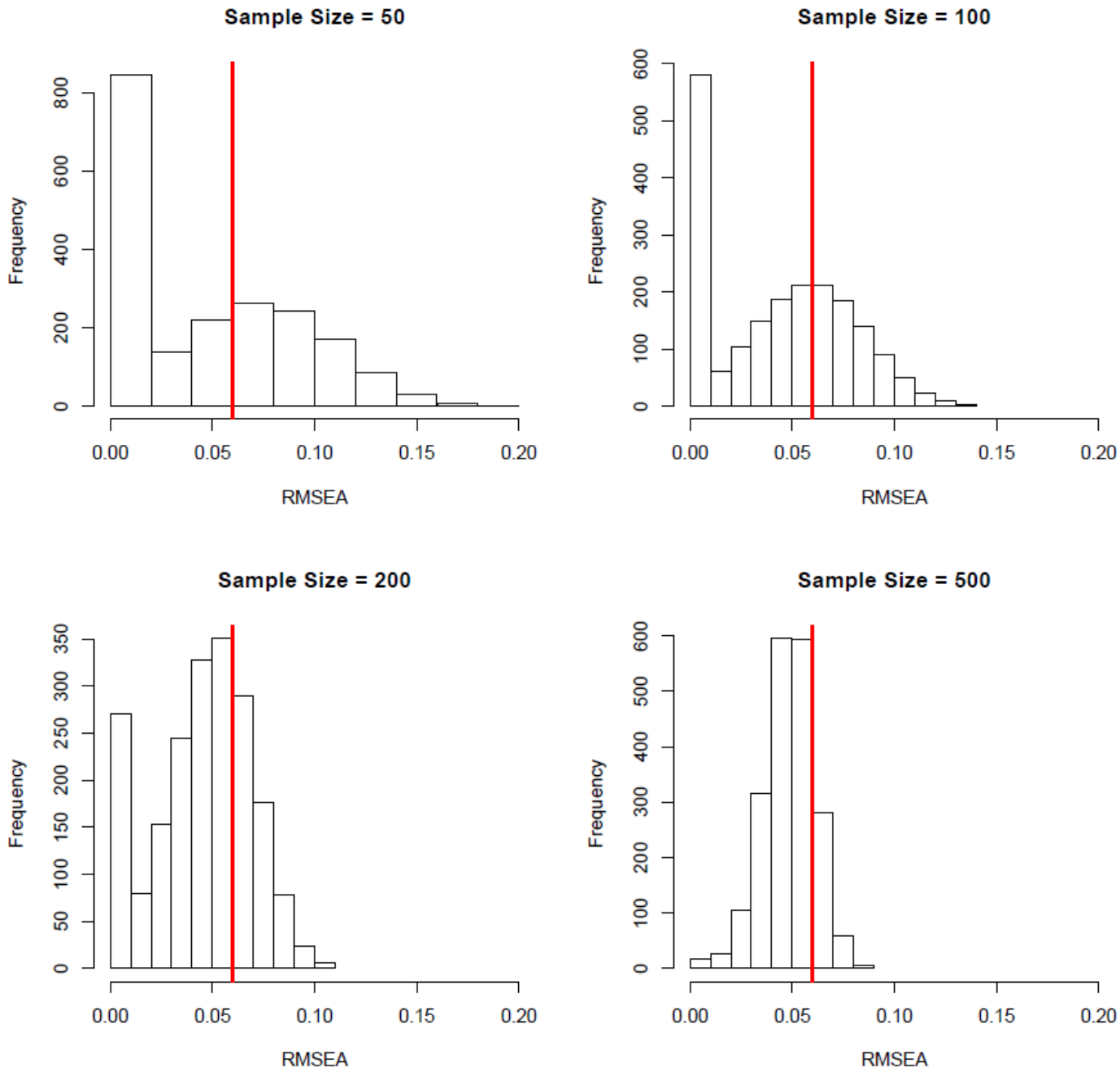
- **การกระจายของข้อมูล**
  - ML จะมีข้อตกลงเบื้องต้นว่าการกระจายของข้อมูลเป็น Multivariate Normality
  - หากไม่เป็นเช่นนั้น ค่า $\chi^2$ จะสูงขึ้นกว่าความเป็นจริง
  - ถึงแม้ Satorra & Bentler (2001) จะนำเสนอวิธีการปรับค่า $\chi^2$ ในกรณีที่การกระจายของข้อมูลไม่เป็น Multivariate Normality การปรับจะปรับเฉพาะการทดสอบ $\chi^2$ ว่าโมเดลมีความเหมาะสมสัมบูรณ์
  - ค่า $\chi^2$ หรือ $F_M$ ที่ใช้ในการคำนวณดัชนีต่างๆ ยังไม่ถูกต้อง

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **การขยายผลการใช้จุดตัด**
  - เนื่องจากดัชนีความเหมาะสม ได้รับผลกระทบจากปัจจัยต่างๆ ในขั้นต้น จุดตัดจึงไม่สามารถใช้ได้กับทุกสถานการณ์
  - เช่น Hu & Bentler (1999) วิเคราะห์จาก Confirmatory Factor Analysis ที่มี 15 ข้อคำถาม ขนาด Standardized factor loadings เท่ากับ .7 ถึง .8

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **วิธีที่ใช้ในการหาจุดตัด**
  - Hu & Bentler (1999) ใช้วิธีในการหาจุดตัด โดยทำให้ค่าผลรวมระหว่าง Type I error และ Type II error มีน้อยที่สุด
  - Type II error จะเกิดได้ ขึ้นอยู่กับว่าผู้วิจัยกำหนดโมเดลที่ต้องการปฏิเสธว่าเป็นอย่างไร
  - นอกจากนี้ วิธีที่ Hu & Bentler (1999) ทำ แตกต่างจากวิธีการทางสถิติโดยทั่วไป ที่กำหนดค่า Type I error และพยายามทำให้ค่า Type II error มีน้อยที่สุด
  - วิธีของ Hu & Bentler (1999) ทำให้เกิดผลแย่คือ กำลังในการปฏิเสธโมเดลที่ไม่ดีน้อยลง หากจำนวนกลุ่มตัวอย่างมากขึ้น


Sample Size = 50
Sample Size = 100
Sample Size = 200
Sample Size = 500
Frequency
RMSEA
0.00
0.05
0.10
0.15
0.20

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **โมเดลที่ต้องการปฏิเสธของแต่ละงานวิจัยแตกต่างกัน**
  - ทำให้ดัชนีความเหมาะสมที่คำนวณได้จากโมเดลเหล่านี้แตกต่างกัน
  - ส่งผลให้งานวิจัยแต่ละชิ้น แนะนำจุดตัดที่แตกต่างกัน

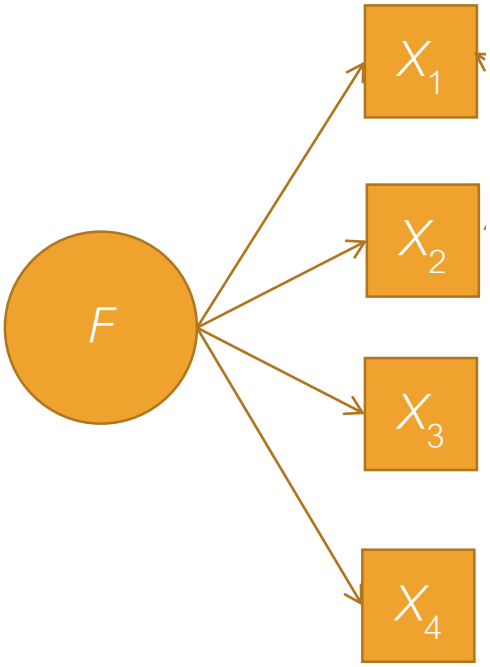
จาก 245 โมเดล ในงานวิจัย 16 ชิ้น

| Statistics | df | ncp | Power | RMSEA | SRMR | CFI | TLI |
|---|---|---|---|---|---|---|---|
| **Across Designs** | | | | | | | |
| **Average** | 132.65 | 31.50 | .421 | .058 | .047 | .955 | .944 |
| **SD** | 202.34 | 57.11 | .343 | .041 | .050 | .056 | .068 |
| **Minimum** | 1.00 | .00 | .050 | .000 | .000 | .635 | .574 |
| **Maximum** | 945.00 | 462.17 | 1.000 | .189 | .312 | 1.000 | 1.000 |

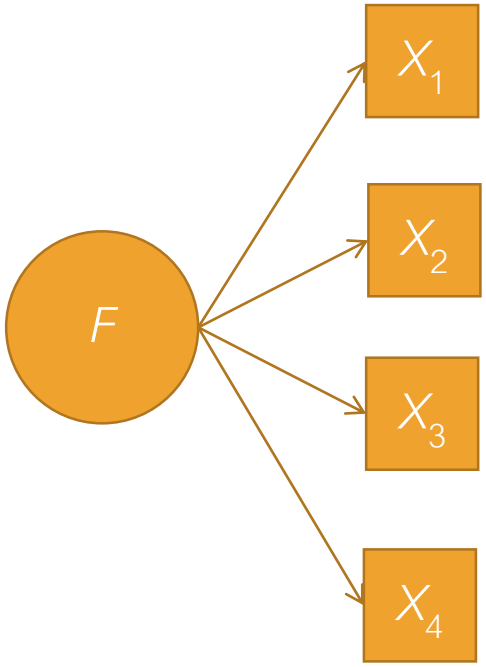
สังเกตว่า ค่าเฉลี่ยของ RMSEA น้อยกว่า .06 ที่ Hu & Bentler แนะนำ

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **งานวิจัยที่หาจุดตัด ไม่ได้คำนึงถึงโมเดลที่มีความเหมาะสมแบบใกล้เคียง**




โมเดลที่ยอมรับได้ในการวิเคราะห์จริง




โมเดลที่ยอมรับได้ในงานวิจัยที่หาจุดตัด

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **ผลลัพธ์ที่แตกต่างกันของดัชนีต่าง ๆ**
  - ดัชนีต่างๆ ให้ผลที่แตกต่างกัน ซึ่งอาจทำให้นักวิจัยตีความตามที่ตนเองต้องการได้

This model did not meet acceptable levels of model fit: $\chi^2$ = 2,389.75, *df* = 344, $p < .0001$, normed $\chi^2$ = 6.9, GFI = .849, SRMR = .134, RMSEA = .075, 90% CI [.072, .078], CFI = .865, and TLI = .852

This model was the best fitting structural model according to multiple GFIs: $\chi^2$ = 1,187.70, *df* = 333, $p < .0001$, normed $\chi^2$ = 3.6, GFI = .923, SRMR = .053, RMSEA = .049, 90% CI [.046, .052], CFI = .944, and TLI = .936.

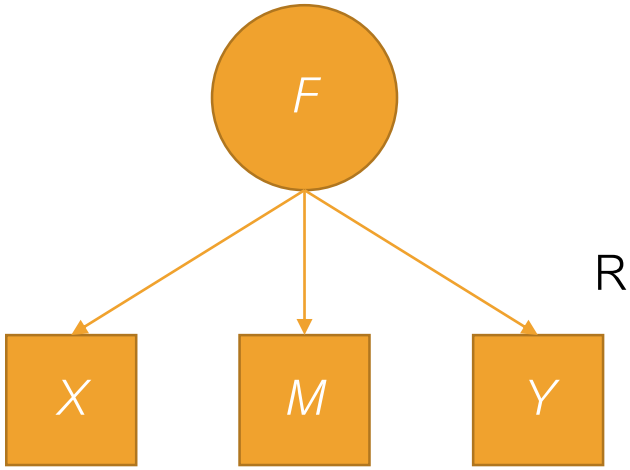
# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **การใส่คำเรียกระดับค่าของดัชนีความเหมาะสม**
  - งานวิจัยต่างๆ ใช้คำเหล่านี้จากค่าที่แตกต่างกัน: “Excellent fit”, “Good fit”, “Adequate fit”
  - เช่น คำว่า “Good fit” เท่ากับ
    - RMSEA = .03 - .05 (Low & Espelage, 2014)
    - RMSEA = .08 (Jensen et al., 2014; Shen et al., 2014)

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **การใส่คำเรียกระดับค่าของดัชนีความเหมาะสม**
  - การใส่คำเรียกระดับต่างๆ ไม่ใช่สิ่งที่ดี
    - งานวิจัยที่มีดัชนีความเหมาะสมที่ดีกว่า อาจไม่ใช่โมเดลที่เหมาะสมเสมอไป

RMSEA > 0




RMSEA = 0

หากผู้วิจัยทราบว่า *X* เกิดก่อน *Y* โมเดลขวาจึงไม่เหมาะสม ทั้งที่มีดัชนีความเหมาะสมดีกว่า

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- **การใส่คำเรียกระดับค่าของดัชนีความเหมาะสม**
  - การใส่คำเรียกระดับต่างๆ ไม่ใช่สิ่งที่ดี
    - ระดับดัชนีความเหมาะสมเดียว อาจแสดงถึงค่าที่ดีในโมเดลหนึ่ง แต่เป็นค่าที่ไม่ดีต่ออีกโมเดลหนึ่ง

RMSEA = .05

**3 ข้อคำถามต่อองค์ประกอบ**

ความสัมพันธ์ระหว่างองค์ประกอบ น้อยกว่า .2

**10 ข้อคำถามต่อองค์ประกอบ**

ความสัมพันธ์ระหว่างองค์ประกอบ มากกว่า .2 มาก

# ปัญหาของการใช้จุดตัดของดัชนีความเหมาะสม

- จากปัญหาที่กล่าวมาข้างต้น เราคิดว่า การหาความเหมาะสมของโมเดลที่ใช้ในปัจจุบันนั้น
  - ไม่ได้ทำให้การวิเคราะห์อยู่ในเชิงปรนัย งานวิจัยไม่สามารถทำซ้ำได้
  - ผลที่ได้ อาจทำให้ได้ผลการสรุปที่ผิด
- แล้วสิ่งที่ต้องการคืออะไร

# วิธีการวัดความเหมาะสมของโมเดลในอุดมคติ

- คัดเลือกโมเดลที่ความเหมาะสมใกล้เคียงกับประชากร และคัดโมเดลที่ความเหมาะสมไม่ใกล้เคียงกับประชากรออก
- สามารถคัดเลือกได้ถูกต้องโดยไม่รับผลกระทบจากปัจจัยในข้างต้น
- ควบคุม Type I error (โอกาสของความผิดพลาดที่เกิดจากการปฏิเสธโมเดลที่ใกล้เคียง)
- กำลังในการปฏิเสธโมเดลที่ไม่ดีสูงขึ้น หากจำนวนกลุ่มัวอย่างสูงขึ้น
- มีวิธีการรวบรวมข้อมูลจากดัชนีหลายชนิด

# วิธีการอื่นที่ถูกพัฒนาขึ้น

- การทดสอบโมเดลที่เหมาะสมใกล้เคียง และเหมาะสมไม่ใกล้เคียง (Test of close fit and not close fit; Browne & Cudeck, 1992)
- การใช้ดัชนีการปรับโมเดลและกำลัง (Modification indices and power approach; Satorra et al., 2009)
- การวิเคราะห์แบบ Bayes (Bayesian analysis; Muthen & Asparouhov, 2012)
- การใช้การจำลองสถานการณ์ (Simulation approach; Millsap, 2012)

# การทดสอบโมเดลที่เหมาะสมใกล้เคียง และเหมาะสมไม่ใกล้เคียง

- ทดสอบความเหมาะสมในภาพรวม
- ทดสอบสมมติฐานดังต่อไปนี้
  - การทดสอบโมเดลที่เหมาะสมใกล้เคียง. $H_0: \varepsilon \leq .05$
  - การทดสอบโมเดลที่เหมาะสมไม่ใกล้เคียง. $H_0: \varepsilon \geq .05$

    โดย $\varepsilon$ คือ RMSEA ในประชากร
- ข้อดี: มีวิธีการนิยามโมเดลใกล้เคียง และจัดการความผิดพลาดจากการสุ่มได้ดี
- ข้อเสีย: ใช้ดัชนีความเหมาะสมในการนิยามโมเดลใกล้เคียง

# การใช้ดัชนีการปรับโมเดลและกำลัง

- ทดสอบความเหมาะสมแยกย่อย
- ตรวจสอบว่าค่า parameter ที่จำกัดไว้นั้นผิดพลาดมากกว่าความผิดพลาดสูงสุดที่รับได้หรือไม่
- ข้อดี: นิยามโมเดลใกล้เคียงโดยใช้พารามิเตอร์
- ข้อเสีย: จัดการความผิดพลาดจากการสุ่มได้ผิดพลาด และสามารถนิยามความผิดพลาดสูงสุดได้จาก parameter ที่จำกัดเท่านั้น

# การวิเคราะห์แบบ Bayes

- ทดสอบความเหมาะสมในภาพรวมและแยกย่อยได้
- ทดสอบความเหมาะสมในภาพรวมด้วย Posterior Predictive P-value (PPP)
- ทดสอบความเหมาะสมแยกย่อยโดยสังเกตว่าช่วงเชื่อถือได้ (credible interval) ของค่า Parameter ที่ไม่สำคัญคลุม 0 หรือไม่
- ข้อดี: ค่า Parameter ที่ไม่สำคัญถูกประมาณค่าได้ โดยไม่จำเป็นต้องกำหนดให้มีค่าเท่ากับ 0
- ข้อเสีย: PPP ได้รับผลกระทบจากปัจจัยหลายอย่างเช่นเดียวกับดัชนีความเหมาะสม การทดสอบแยกย่อยได้รับผลกระทบจากจำนวนกลุ่มตัวอย่างและ priors นอกจากนี้ ยังไม่มีจุดตัดที่ชัดเจนสำหรับ PPP

# การใช้การจำลองสถานการณ์

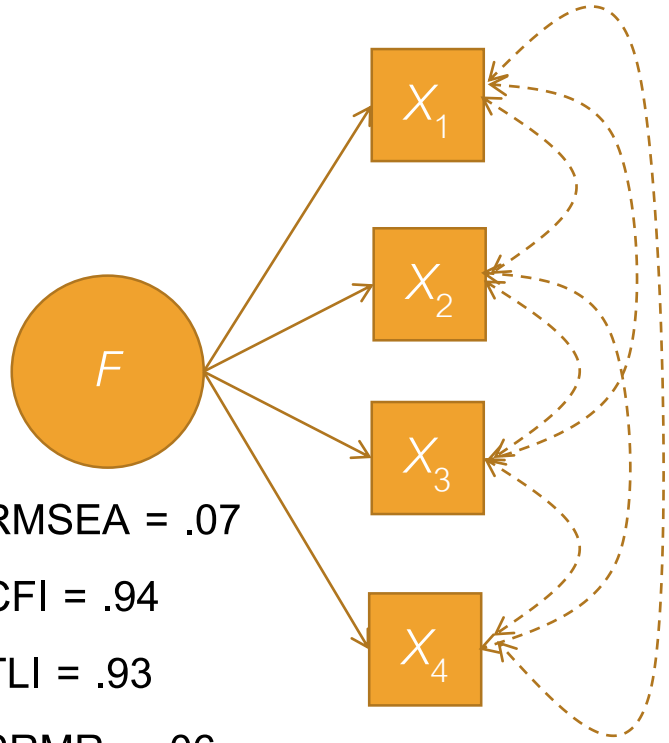
- ทดสอบความเหมาะสมในภาพรวม
- นำโมเดลที่วิเคราะห์ได้ ไปเพิ่มค่า Parameter ที่ไม่สำคัญ
- นำโมเดลที่ดัดแปลงนี้ ไปสร้างข้อมูลจำนวนมาก และทดสอบว่าค่าดัชนีเหมาะสมจะมีการกระจายแบบใด
- ข้อดี: นิยามโมเดลใกล้เคียงในระดับ Parameter และสามารถใช้ได้กับดัชนีเหมาะสมทุกประเภท
- ข้อเสีย: การตัดสินว่า Null hypothesis “ถูกต้อง” และถูกจำกัดให้เพิ่มค่า Parameter ที่ไม่สำคัญได้เพียงแค่ 1 ชุด (หรือมากกว่านั้นไม่มาก)

# วิธีการแบบบูรณาการ

- วิธีการแบบบูรณาการจึงถูกพัฒนาขึ้นมา
  - เพื่อรวบรวมข้อดีของแต่ละวิธีอื่นที่ถูกพัฒนาขึ้น
  - เพื่อแก้ไขข้อเสียของวิธีเหล่านั้น
- แบ่งเป็น 2 ส่วน
  - การทดสอบในภาพรวม ด้วยดัชนีความเหมาะสม
  - การทดสอบแยกย่อย ด้วยค่า Parameter ที่คาดว่าจะเปลี่ยนแปลง (Expected Parameter Changes; EPC)

# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

## โมเดลความเหมาะสมใกล้เคียง

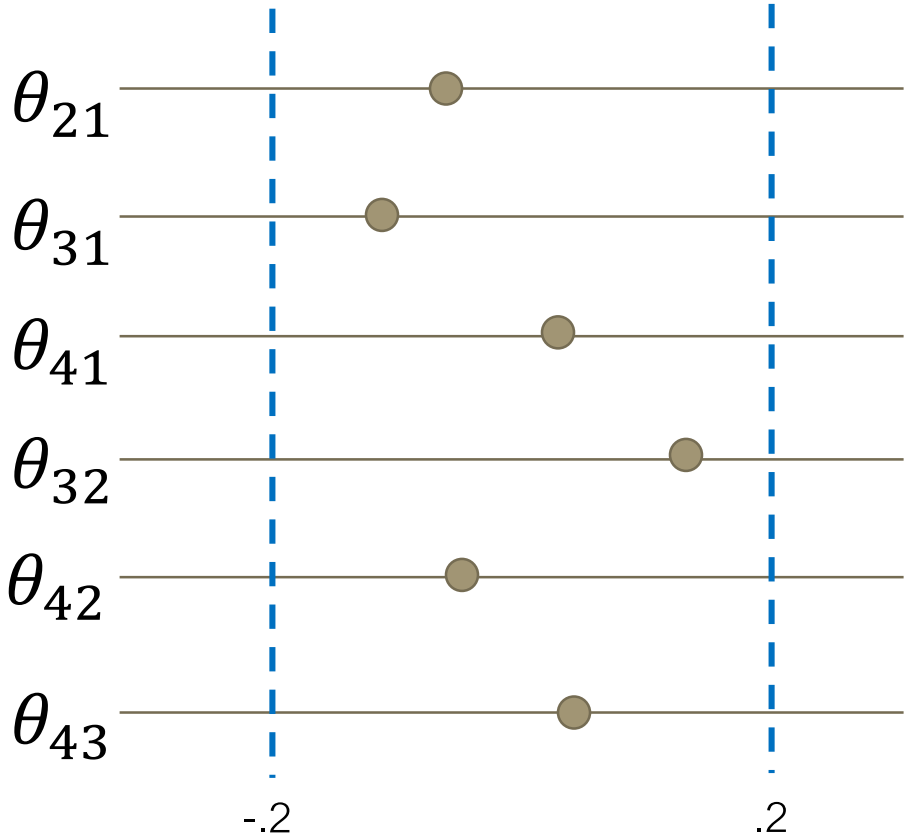



RMSEA = .07

CFI = .94

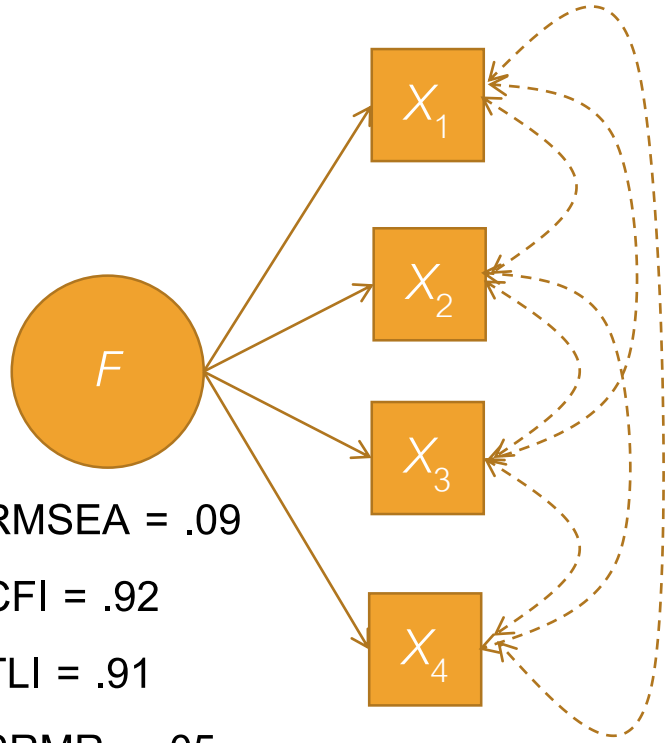
TLI = .93

SRMR = .06



# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

## โมเดลความเหมาะสมใกล้เคียง

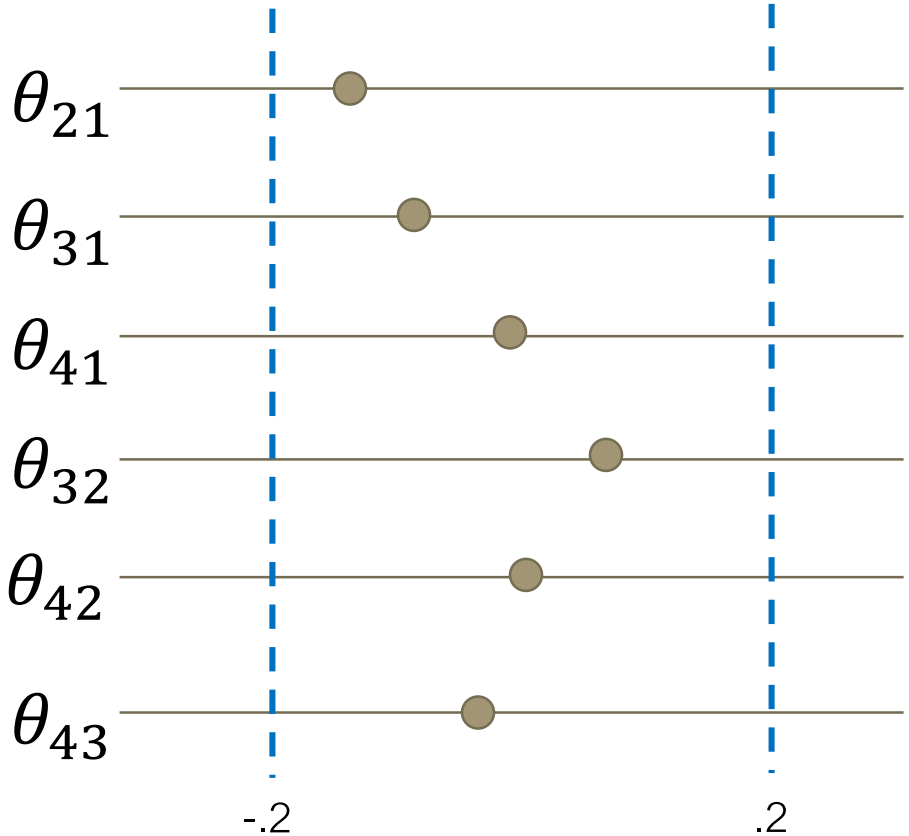



RMSEA = .09

CFI = .92

TLI = .91

SRMR = .05



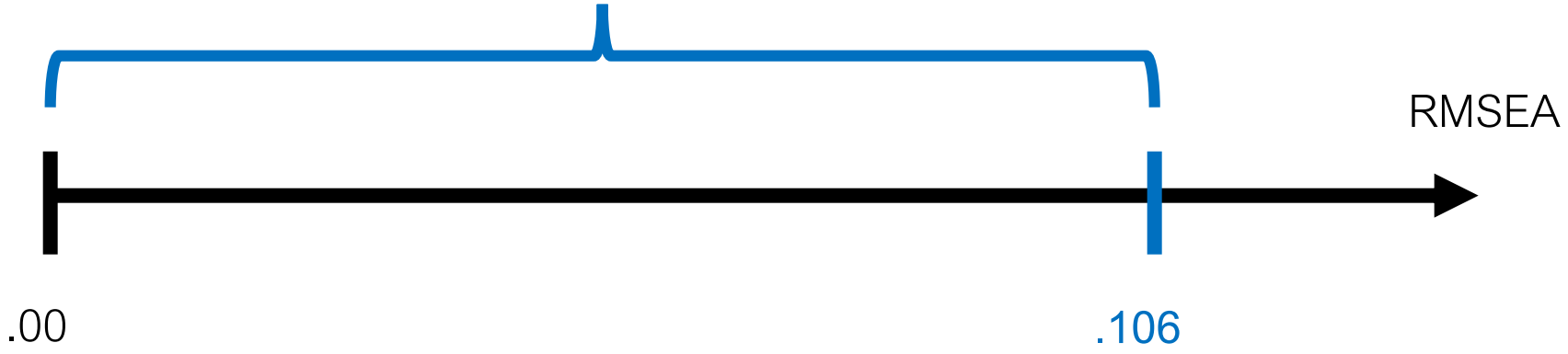
# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

นำโมเดลที่มีค่าความเหมาะสมแย่ที่สุดออกมา

สร้างข้อมูลจำนวนมากออกมา หาค่าดัชนีเหมาะสมในกลุ่มตัวอย่าง

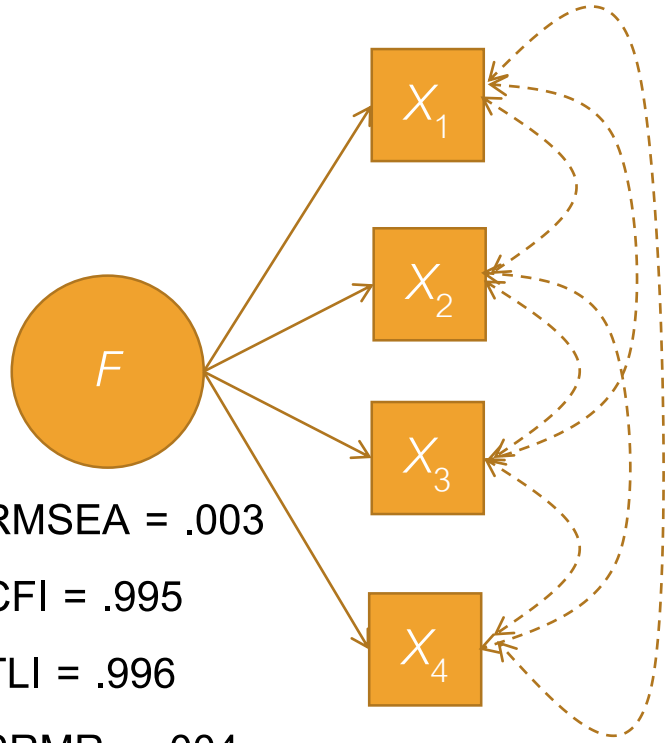
หา Critical Value

ดัชนีเหมาะสม สำหรับโมเดลใกล้เคียง



# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

โมเดลความเหมาะสมไม่ใกล้เคียง

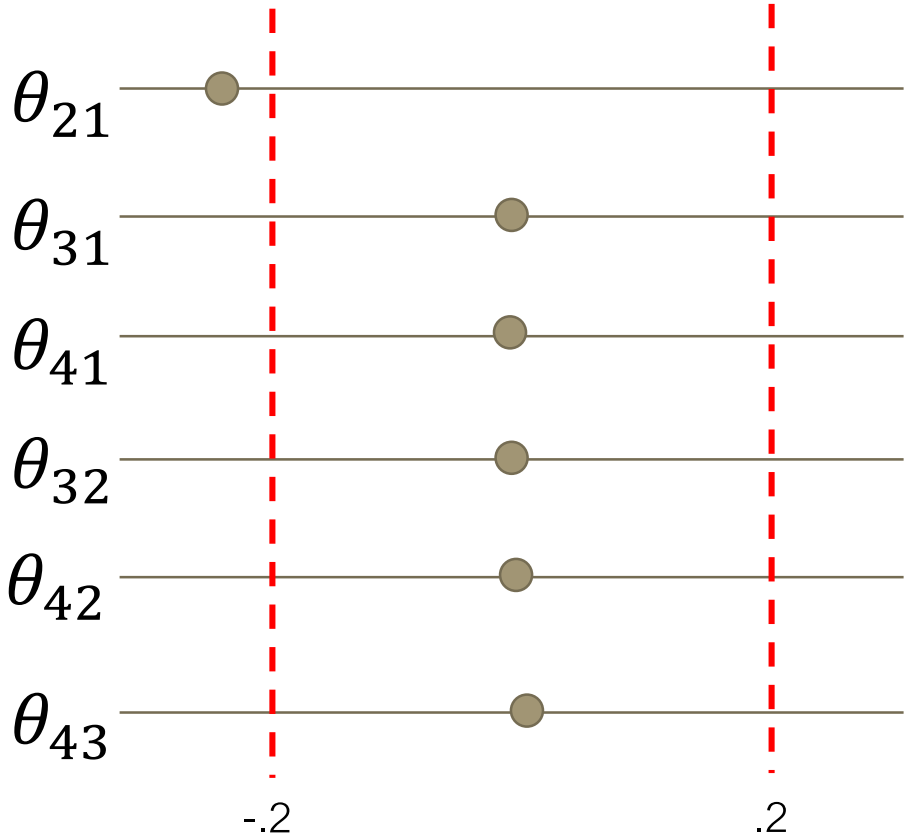



RMSEA = .003

CFI = .995

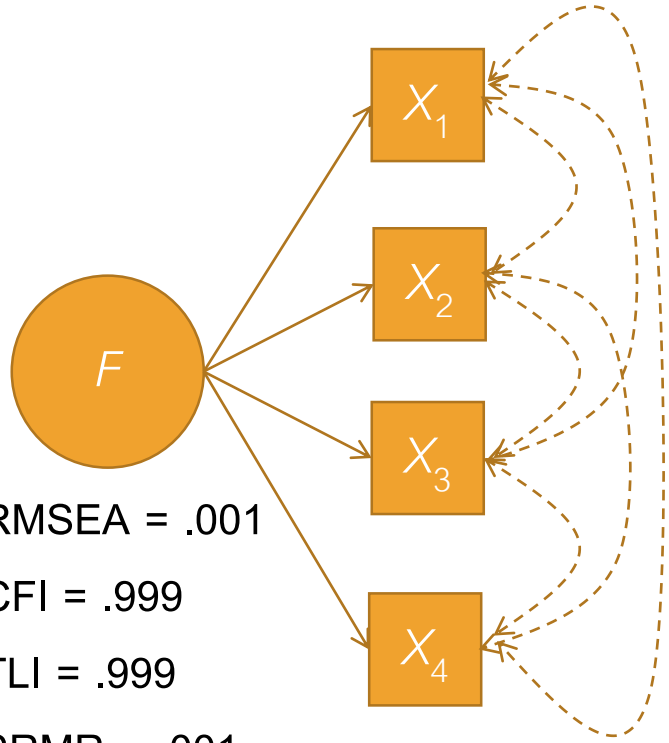
TLI = .996

SRMR = .004



# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

โมเดลความเหมาะสมไม่ใกล้เคียง

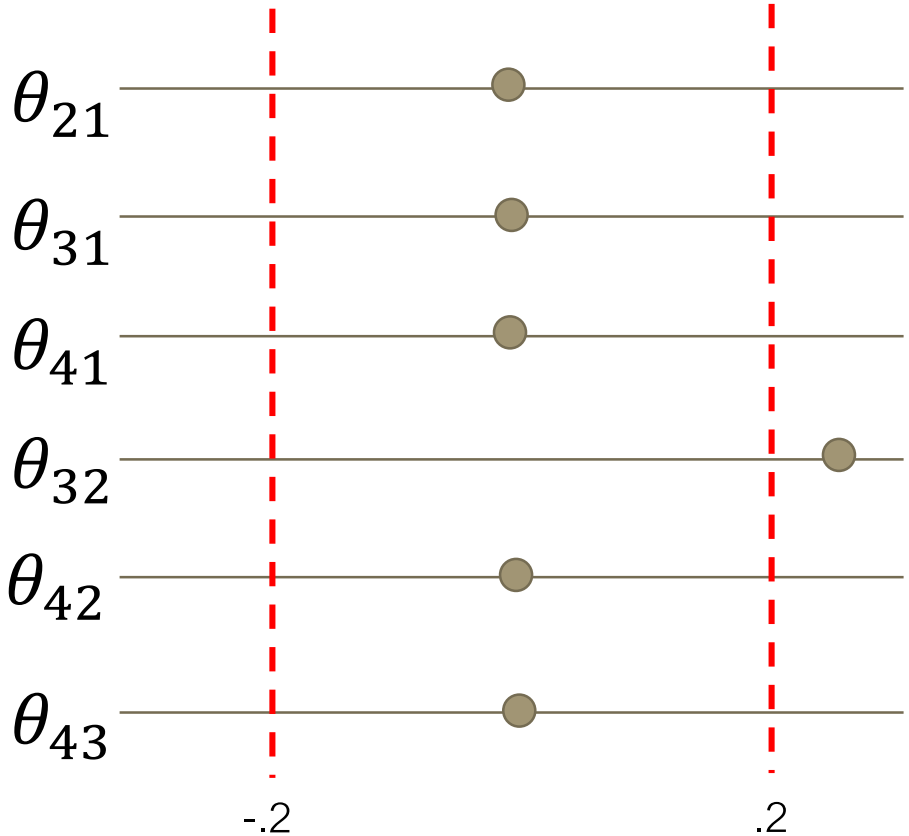



RMSEA = .001

CFI = .999

TLI = .999

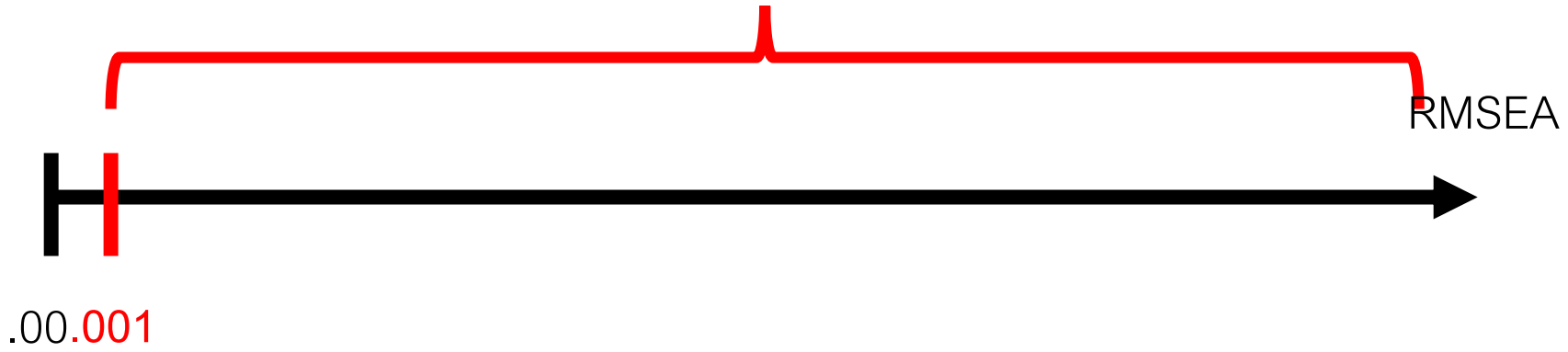
SRMR = .001



# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

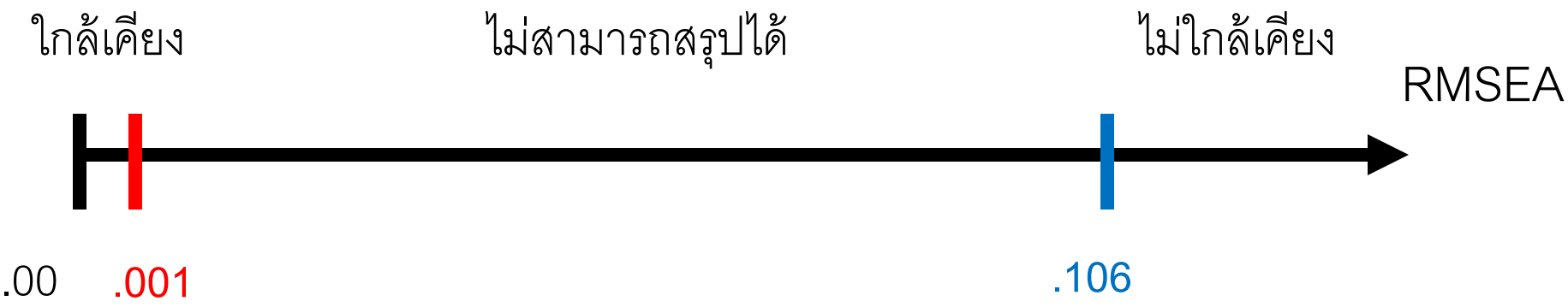
นำโมเดลที่มีค่าความเหมาะสมดีที่สุดออกมา → สร้างข้อมูลจำนวนมากออกมา หาค่าดัชนีเหมาะสมในกลุ่มตัวอย่าง → หา Critical Value

ดัชนีเหมาะสม สำหรับโมเดลไม่ใกล้เคียง



# วิธีการแบบบูรณาการ: ความเหมาะสมในภาพรวม

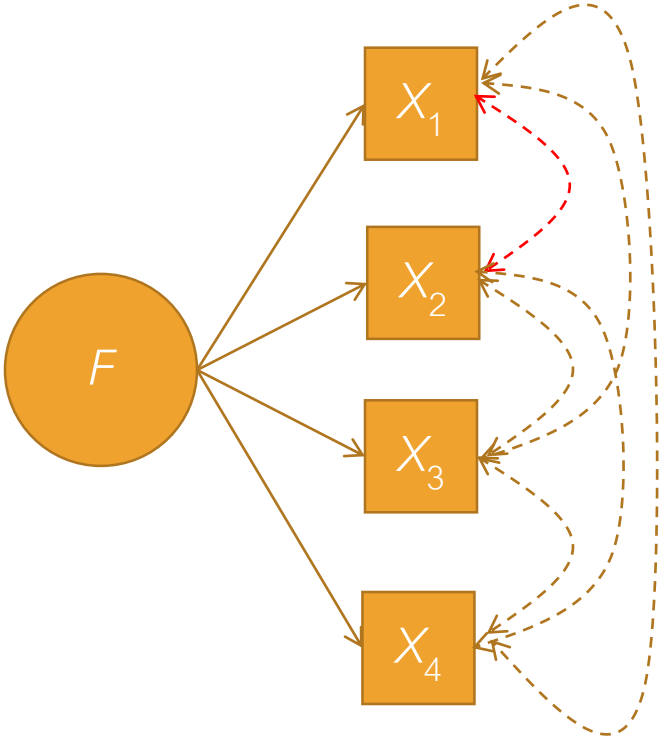
นำค่าดัชนีเหมาะสมจากการวิเคราะห์จริงมาเปรียบเทียบกับสเกล




ทำกระบวนการนี้ กับดัชนีอื่นๆ ผลที่ได้ทั้งหมดจะนำมารวมกัน

# วิธีการแบบบูรณาการ: ความเหมาะสมแยกย่อย

- ใช้ดัชนีการปรับโมเดล (Modification Indices) และค่า Parameter ที่คาดว่าจะเปลี่ยนแปลง (Expected parameter changes)
- นำช่วงเชื่อมั่นของค่า Parameter ที่คาดว่าจะปลี่ยนแปลงไปเปรียบเทียบกับช่วงใกล้เคียง


F
X1
X2
X3
X4

# วิธีการแบบบูรณาการ: ความเหมาะสมแยกย่อย

- นำผลการเปรียบเทียบได้ทั้งหมดมารวมกัน
  - ถ้าช่วงเชื่อมั่นอันใดอันหนึ่งแสดงว่าการวิเคราะห์มีกำลังไม่พอแล้ว การวิเคราะห์กำลังไม่พอ (Underpowered)
  - ถ้าช่วงเชื่อมั่นอันใดอันหนึ่งแสดงว่าโมเดลไม่ใกล้เคียงแล้ว โมเดลไม่ใกล้เคียง (Severe misspecification)
  - ถ้าช่วงเชื่อมั่นอันใดอันหนึ่งแสดงว่าไม่สามารถสรุปได้แล้ว ไม่สามารถสรุปได้ว่าโมเดลใกล้เคียงหรือไม่ใกล้เคียง (Inconclusive)
  - นอกเหนือจากข้างต้น โมเดลใกล้เคียง (Trivial misspecification)

# การศึกษาการจำลองสถานการณ์

- การศึกษาที่ 1: CFA แบบมีสององค์ประกอบ
- การศึกษาที่ 2: Latent growth curve

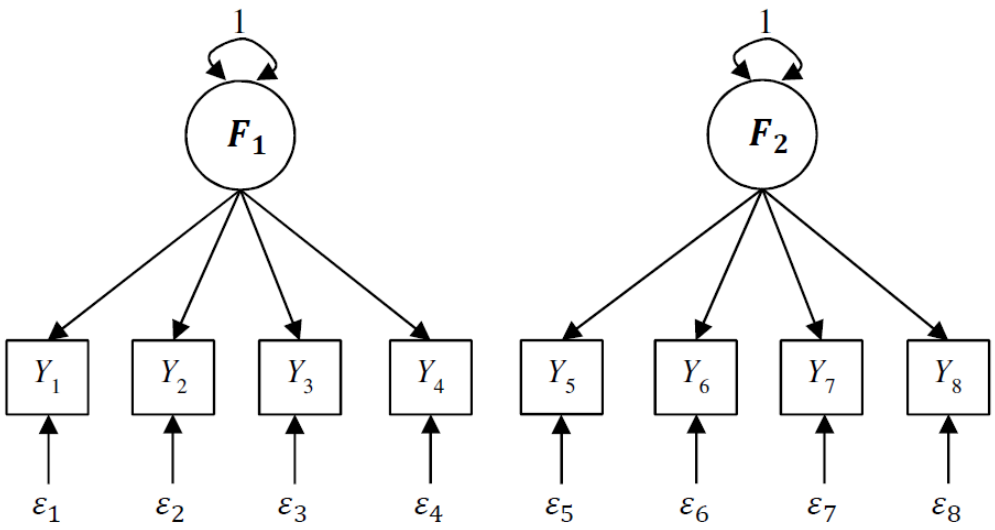
# การศึกษาการจำลองสถานการณ์

- ทดสอบผลของวิธีแบบบูรณาการ (การศึกษาที่ 1 และ 2)
- เปรียบเทียบผลของวิธีแบบบูรณาการกับวิธีอื่นๆ (การศึกษาที่ 1)

# การศึกษาที่ 1 (การออกแบบงานวิจัย)

## เงื่อนไขที่ใช้ศึกษา

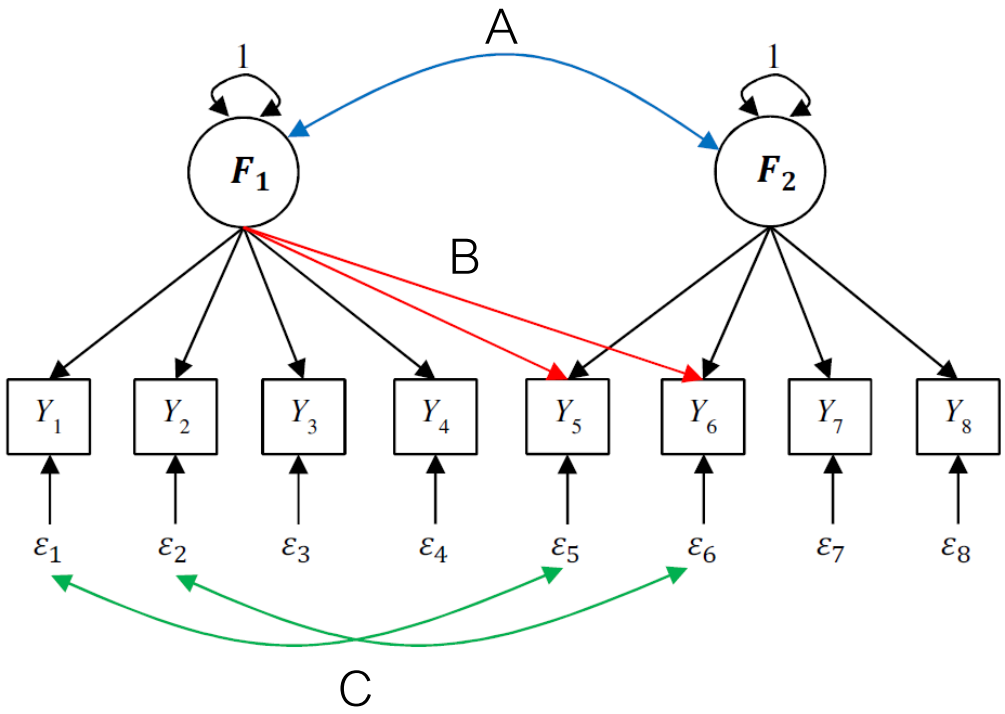
- 10 รูปแบบของโมเดลในประชากรที่ไม่ตรงกับการวิเคราะห์
- Standardized Factor Loadings: .5, .7
- จำนวนข้อคำถาม: 8, 16
- ขนาดกลุ่มตัวอย่าง: 125, 250, 500, 1000, 2000, 4000
- ระดับที่กำหนดว่าโมเดลมีความเหมาะสมใกล้เคียง: .1, .3



# การศึกษาที่ 1 (การออกแบบงานวิจัย)

รูปแบบของโมเดลในประชากรที่ไม่ตรงกับการวิเคราะห์

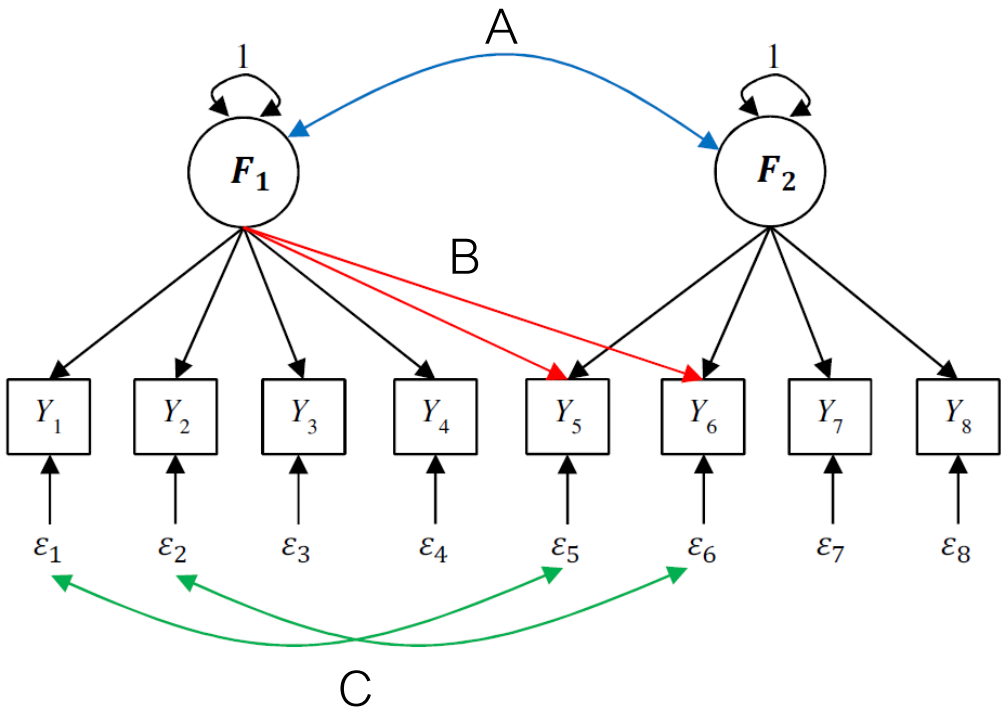
| Types | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| A | 1 | 2 | 3 |
| B | 4 | 5 | 6 |
| C | 7 | 8 | 9 |
| No Misspecification | | | 10 |



# การศึกษาที่ 1 (การออกแบบงานวิจัย)

รูปแบบของโมเดลในประชากรที่ไม่ตรงกับการวิเคราะห์

| Types | Level 1 | Level 2 | Level 3 |
|---|---|---|---|
| A | .1 | .3 | .9 |
| B | .1 | .3 | .9 |
| C | .1 | .3 | .9 |
| No Misspecification | | | .0 |



# การศึกษาที่ 1 (การออกแบบงานวิจัย)

- ใช้ดัชนีความเหมาะสม 4 แบบ: RMSEA, CFI, TLI, and SRMR
- การกำหนดระดับความใกล้เคียงของโมเดลในวิธีแบบบูรณาการ
  - Factor correlation อยู่ในช่วง -.1 และ .1 (หรือ -.3 และ .3)
  - Standardized cross loadings อยู่ในช่วง -.1 และ .1 (หรือ -.3 และ .3)
  - Measurement error correlations อยู่ในช่วง -.1 และ .1 (หรือ -.3 และ .3)

# การศึกษาที่ 1 (การออกแบบงานวิจัย)

## วิธีอื่นที่ใช้ในการทดสอบความเหมาะสมของโมเดล

- จุดตัดดัชนีความเหมาะสม ตามคำแนะนำของ Hu & Bentler (1999)
- การทดสอบโมเดลที่เหมาะสมใกล้เคียง และเหมาะสมไม่ใกล้เคียง (Test of close fit and not close fit)
- การใช้ดัชนีการปรับโมเดลและกำลัง (Modification indices and power approach)
- การวิเคราะห์แบบ Bayes (Bayesian analysis)
- การใช้การจำลองสถานการณ์ (Simulation approach)

# การศึกษาที่ 1 (การออกแบบงานวิจัย)

- ผลที่ต้องการทดสอบ
  - อัตราส่วนที่พบผลที่ไม่สามารถสรุปได้
  - อัตราส่วนของการปฏิเสธโมเดล ระหว่างผลที่สามารถสรุปได้
- ใช้ Eta-squares ในการทดสอบอิทธิพลของเงื่อนไขต่างๆ

# การศึกษาที่ 1 (ผลการจำลองสถานการณ์)

- อัตราส่วนในการปฏิเสธโมเดล จากวิธีการแบบบูรณาการ

ค่าของ Parameter ที่ไม่ได้กำหนดในการวิเคราะห์

| | | 0 | .1 | .3 | .9 |
|---|---|---|---|---|---|
| ระดับที่กำหนดว่าโมเดลมีความเหมาะสมใกล้เคียง | .1 | .00 | .74 | 1.00 | 1.00 |
| | .3 | .00 | .00 | .66 | 1.00 |

ผลปฏิสัมพันธ์นี้ไม่เกิดขึ้นในวิธีการทดสอบอื่น

# การศึกษาที่ 1 (ผลการจำลองสถานการณ์)

- อัตราส่วนในการปฏิเสธโมเดลจากวิธีการแบบบูรณาการไม่ได้รับอิทธิพลจาก
  - รูปแบบของ Parameter ที่ไม่ได้กำหนด
  - ขนาดกลุ่มตัวอย่าง
  - ขนาดของ Standardized factor loadings
  - จำนวนข้อคำถาม
- วิธีการทดสอบอื่นจะได้รับอิทธิพลจากปัจจัยอย่างใดอย่างหนึ่ง

# การศึกษาที่ 1 (ผลการจำลองสถานการณ์)

- อัตราส่วนของการเจอผลที่ไม่สามารถสรุปได้ ของวิธีการแบบบูรณาการ

ขนาดกลุ่มตัวอย่าง

| 125 | 250 | 500 | 1000 | 2000 | 4000 |
|---|---|---|---|---|---|
| .74 | .65 | .43 | .36 | .29 | .28 |

# การศึกษาที่ 1 (ผลการจำลองสถานการณ์)

- อัตราส่วนของการเจอผลที่ไม่สามารถสรุปได้ ของวิธีการแบบบูรณาการ

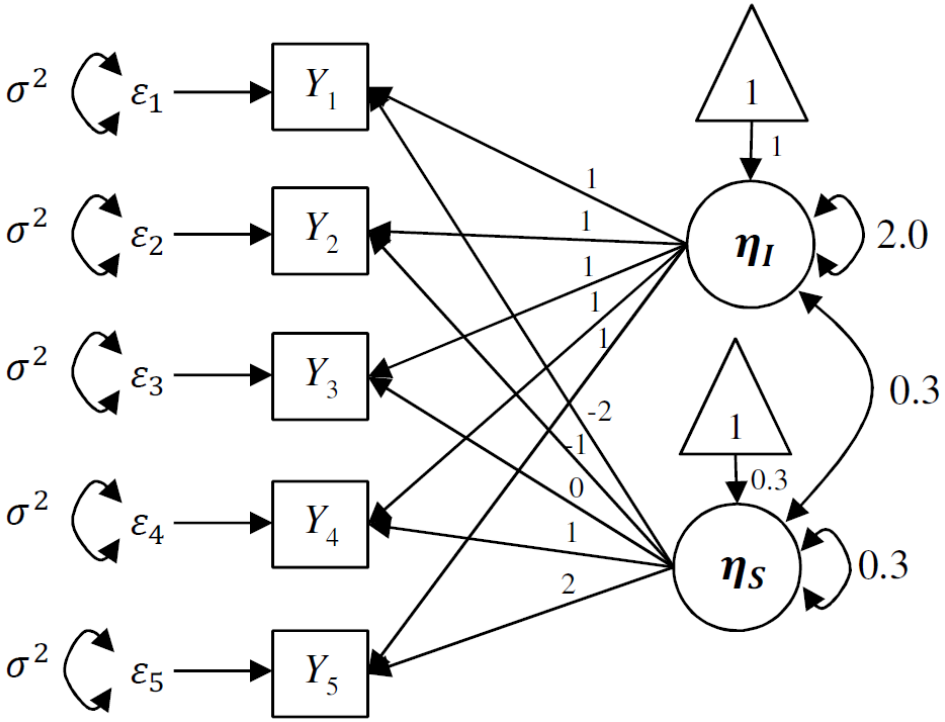
ค่าของ Parameter ที่ไม่ได้กำหนดในการวิเคราะห์

| | | 0 | .1 | .3 | .9 |
|---|---|---|---|---|---|
| ระดับที่กำหนดว่าโมเดลมีความเหมาะสมใกล้เคียง | .1 | .84 | .95 | .46 | .01 |
| | .3 | .34 | .35 | .87 | .10 |

# การศึกษาที่ 2 (การออกแบบงานวิจัย)

## เงื่อนไขที่ใช้ศึกษา

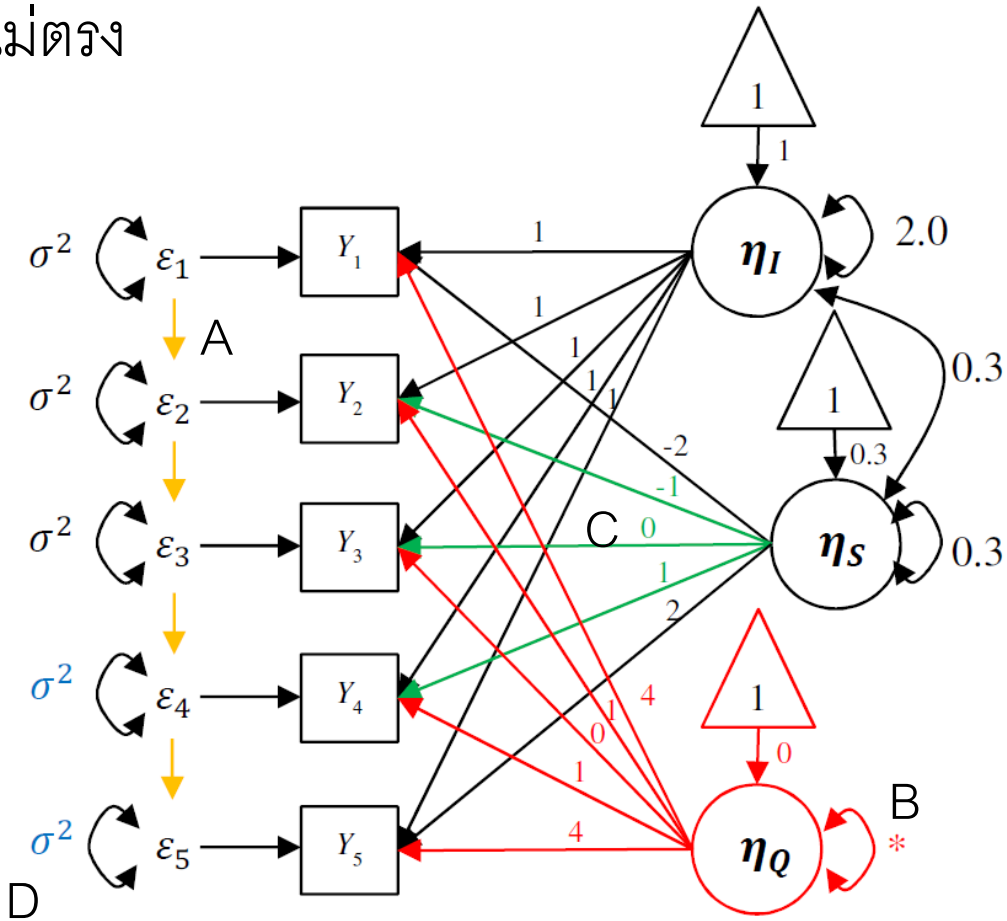
- 13 รูปแบบของโมเดลในประชากรที่ไม่ตรงกับการวิเคราะห์
- ขนาดของความแปรปรวนของค่าคงเหลือในการวัดครั้งที่ 1: 0.5, 2, 3.5
- ขนาดกลุ่มตัวอย่าง: 125, 250, 500, 1000



# การศึกษาที่ 2 (การออกแบบงานวิจัย)

รูปแบบของโมเดลในประชากรที่ไม่ตรงกับการวิเคราะห์

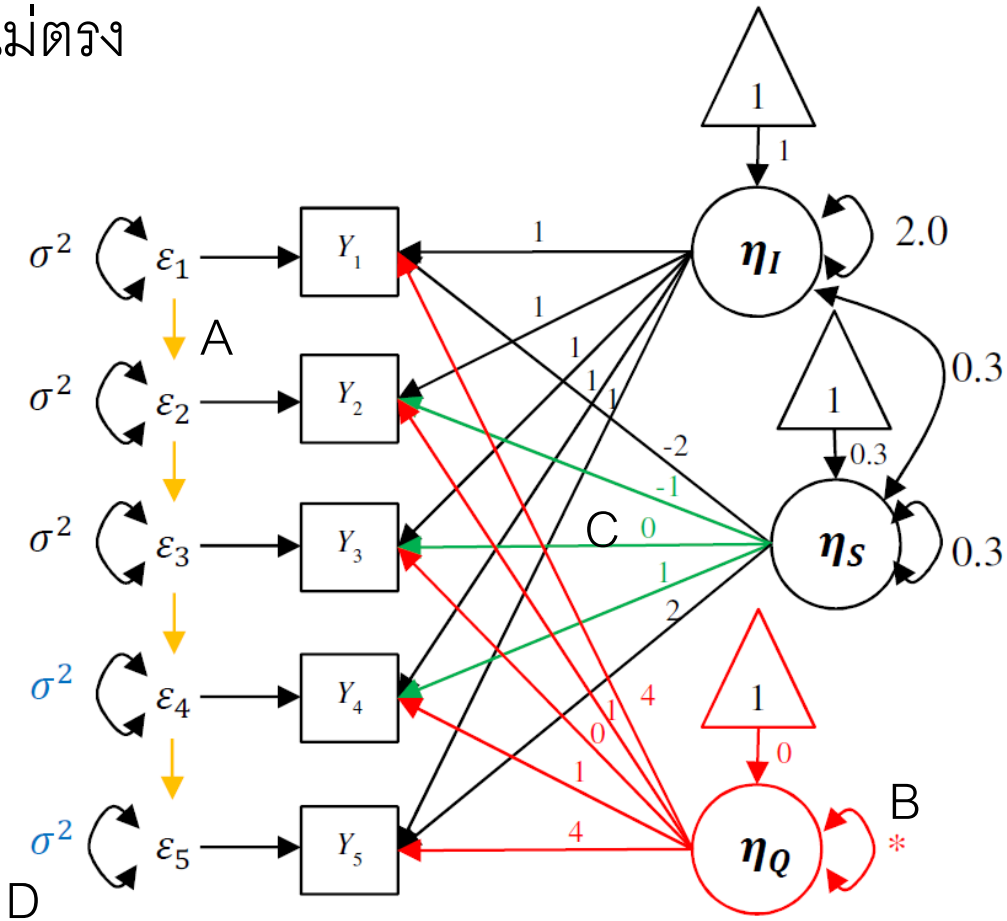
| Types | Trivial | Severe | Very Severe |
|---|---|---|---|
| A | 1 | 2 | 3 |
| B | 4 | 5 | 6 |
| C | 7 | 8 | 9 |
| D | 10 | 11 | 12 |
| No Misspecification | | | 13 |



# การศึกษาที่ 2 (การออกแบบงานวิจัย)

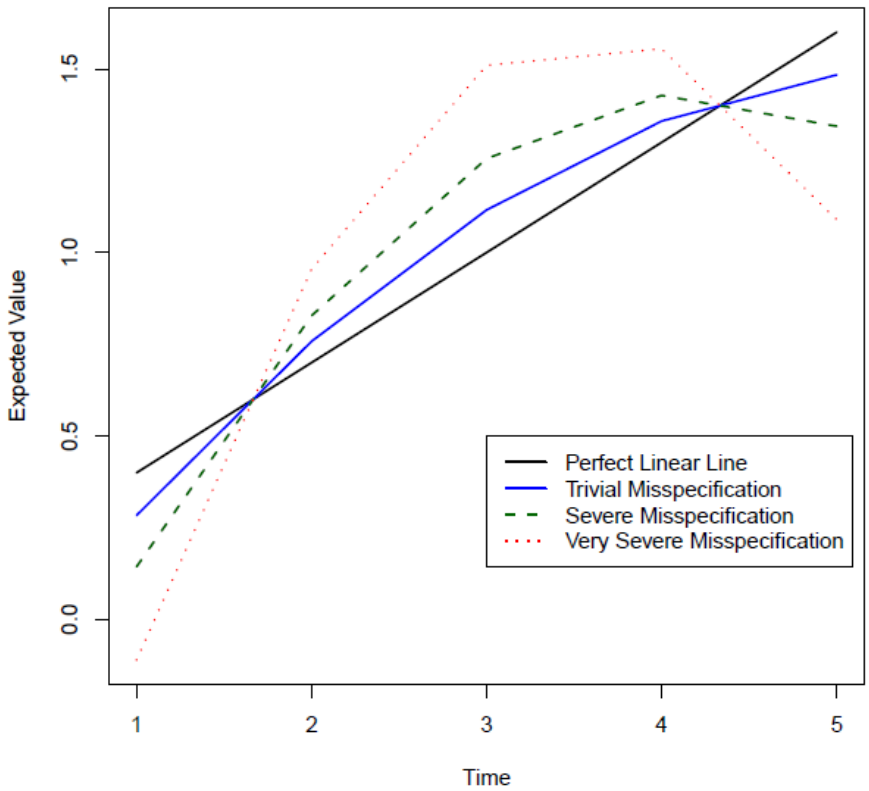
รูปแบบของโมเดลในประชากรที่ไม่ตรงกับการวิเคราะห์

| Types | Trivial | Severe | Very Severe |
|---|---|---|---|
| A | .1 | .3 | .7 |
| B | 0.0034 | 0.0164 | 0.0650 |
| C | -2, -0.67, 0.37, 1.98, 2 | -2, -0.22, 0.88, 3.31, 2 | -2, 0.83, 2.08, 6.45, 2 |
| D | 1.376 | 2.78 | 23.84 |
| No Misspecification | | | N/A |

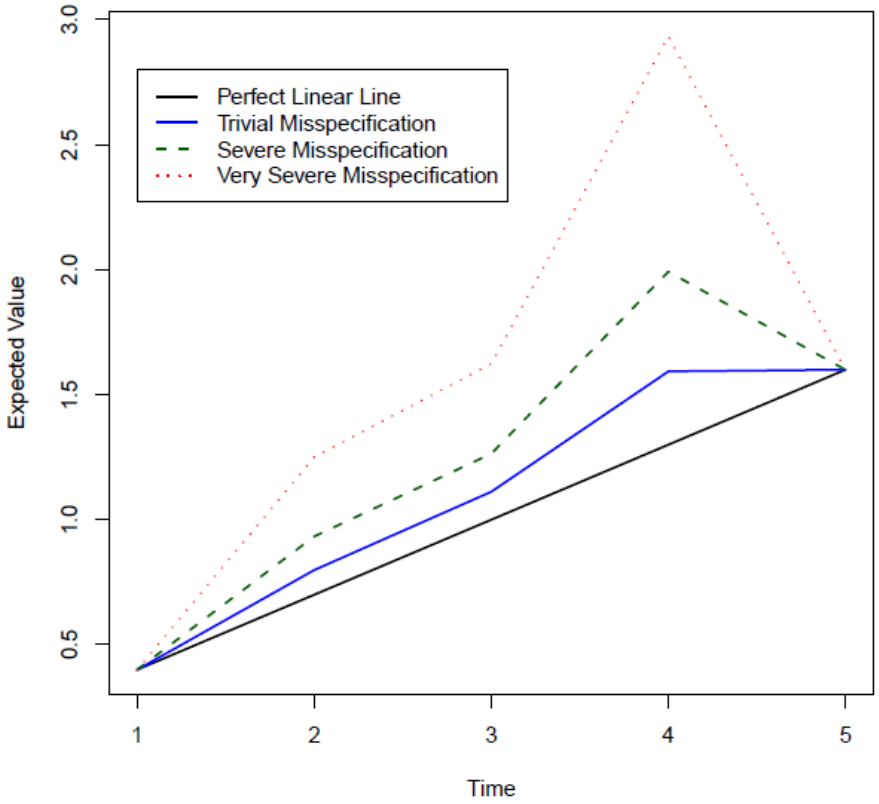


# การศึกษาที่ 2 (การออกแบบงานวิจัย)

Type B Misspecification




Type C Misspecification



# การศึกษาที่ 2 (การออกแบบงานวิจัย)

รายละเอียดของวิธีการแบบบูรณาการ

- ใช้ดัชนีความเหมาะสม 8 แบบ: RMSEA, CFI, TLI, SRMR, marginal pseudo $R^2$, conditional pseudo $R^2$, conditional concordance correlation, and average concordance correlation

# การศึกษาที่ 2 (การออกแบบงานวิจัย)

- การกำหนดระดับความใกล้เคียงของโมเดลในวิธีแบบบูรณาการ
  - Residual error correlation อยู่ในช่วง .2 or -.2
  - สัดส่วนระหว่าง Residual variances สูงสุดและต่ำที่สุด ไม่เกิน 1.926
  - ค่า Factor loadings ขององค์ประกอบการเปลี่ยนแปลงแบบเส้นตรงสามารถเบี่ยงเบนได้ระดับหนึ่ง โดยที่ $R^2$ ของ Marginal means ที่เบี่ยงเบนถูกอธิบายและ Marginal means ดั้งเดิม ไม่น้อยกว่า .9
  - องค์ประกอบการเปลี่ยนแปลงแบบเส้นโค้งมีความแปรปรวนไม่เกิน 0.083

# การศึกษาที่ 2 (ผลการจำลองสถานการณ์)

- อัตราส่วนในการปฏิเสธโมเดล จากวิธีการแบบบูรณาการ

ค่าของ Parameter ที่ไม่ได้กำหนดในการวิเคราะห์

| | L0 | L1 | L2 | L3 |
|---|---|---|---|---|
| ขนาดกลุ่มตัวอย่าง 125 | NA | NA | 1.00 | 1.00 |
| 250 | NA | NA | 1.00 | 1.00 |
| 500 | .00 | .33 | .91 | 1.00 |
| 1000 | .00 | .09 | .83 | 1.00 |

NA = อัตราส่วนของผลที่ไม่สามารถสรุปได้มากกว่า .90

# การศึกษาที่ 2 (ผลการจำลองสถานการณ์)

- อัตราส่วนในการปฏิเสธโมเดลจากวิธีการแบบบูรณาการไม่ได้รับอิทธิพลจาก
  - รูปแบบของ Parameter ที่ไม่ได้กำหนด
  - ขนาดของ Residual variances
- ขนาดของกลุ่มตัวอย่างที่มากขึ้น ทำให้อัตราส่วนในการปฏิเสธโมเดลลดลงเล็กน้อย

# การศึกษาที่ 2 (ผลการจำลองสถานการณ์)

- อัตราส่วนของการเจอผลที่ไม่สามารถสรุปได้ ของวิธีการแบบบูรณาการ

ขนาดกลุ่มตัวอย่าง

| 125 | 250 | 500 | 1000 |
|---|---|---|---|
| .85 | .63 | .48 | .28 |

# การศึกษาที่ 2 (ผลการจำลองสถานการณ์)

- อัตราส่วนของการเจอผลที่ไม่สามารถสรุปได้ ของวิธีการแบบบูรณาการ

ค่าของ Parameter ที่ไม่ได้กำหนดในการวิเคราะห์

| L0 | L1 | L2 | L3 |
|---|---|---|---|
| .69 | .78 | .59 | .25 |

ระดับที่กำหนดว่าโมเดลมีความเหมาะสมใกล้เคียง

# การอภิปรายผล

- วิธีการแบบบูรณาการสามารถแก้ปัญหาของการใช้จุดตัดดัชนีความเหมาะสมได้ทั้งหมด
  - ไม่ได้รับอิทธิพลจากรูปแบบของโมเดลหรือจำนวนกลุ่มตัวอย่าง
  - ทดสอบโดยคำนึงถึงโมเดลที่ความเหมาะสมใกล้เคียง ไม่ใช่โมเดลที่มีความเหมาะสมแบบสัมบูรณ์
  - รวบรวมข้อมูลจากดัชนีความเหมาะสมหลายรูปแบบ
- วิธีการแบบบูรณาการได้ผลลัพธ์ที่ดีกว่าวิธีการอื่นๆ ทั้งหมด

# การอภิปรายผล

- จุดด้อย 3 จุดของวิธีการแบบบูรณาการ
  - ผู้วิจัยสามารถให้คำนิยามโมเดลมีความเหมาะสมใกล้เคียงแตกต่างกัน
  - มีผลที่ไม่สามารถสรุปได้
  - ต้องการกลุ่มตัวอย่างขนาดใหญ่ (> 500) ซึ่งมากกว่าวิธีการอื่นๆ
- นักวิเคราะห์จะต้องระมัดระวังในการใช้จุดตัดของดัชนีความเหมาะสม

# ทิศทางในอนาคต

- การเปรียบเทียบโมเดล ทั้ง Nested Model Comparison และ Nonnested model comparison
- การใช้วิธีการแบบบูรณาการในกรณีที่มีข้อมูลสูญหาย
- การกระจายของข้อมูลไม่เป็นโค้งปกติ

# ขอบคุณครับ

เชิญถามคำถาม หรือให้ข้อเสนอแนะ