



中山大學心理學系

SUN YAT-SEN UNIVERSITY DEPARTMENT OF PSYCHOLOGY

渐近测量不变性中 先验方差的选择研究

报告人：张沥今

指导老师：潘俊豪副教授



01

引言

研究背景与研究意义

02

模拟研究

在不同数据条件下，不同先验分布的表现

03

实证研究

不同性别在社会支持量表中的反应差异

04

总讨论

总结与展望



中山大學心理學系

SUN YAT-SEN UNIVERSITY DEPARTMENT OF PSYCHOLOGY

引言

PART ONE

研究背景与研究意义

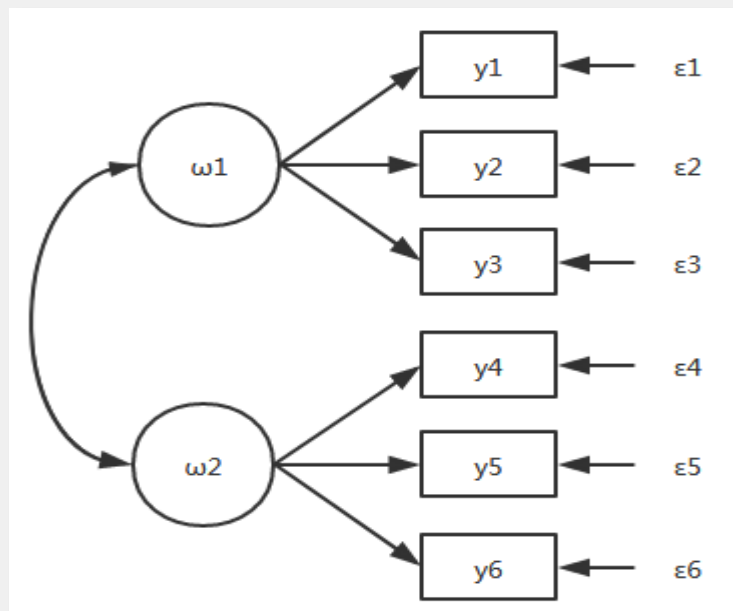


多组比较

- **群体间差异的检验一直受到了广泛的关注**
 - 比较不同性别、年级、文化中群体的特质水平、变量间关系等(欧阳湘子, 田伟, 辛涛, 詹沛达, 2016; Billiet, 2013等)
- **潜变量 (因子)**
 - 抽象的构念(Latent Variables): 如态度、价值观、能力等
 - 不能直接被观测的、而需要用量表条目等可以直接观测的变量来反映的潜变量

潜变量建模

- 验证性因子分析(Confirmatory Factor Analysis, CFA)模型



测量不变性

模型		定义
1	形态不变性	各组具有相同的因子结构
2	载荷不变性	条目载荷跨组相等
3	截距不变性	条目截距跨组相等
4	误差方差不变性	条目误差方差跨组相等

- 多组验证性因子分析(Multi-group Confirmatory Factor Analysis)
- 在满足截距不变性的基础上可以进行结构不变性的检验，包括因子均值差异、因子方差/协方差差异的检验

多组验证性因子分析

- **限制过于严格：载荷和截距参数严格跨组不变**
 - 模型很容易被拒绝(Asparouhov & Muthén, 2014; Kim et al., 2017)
 - 在实证研究中，截距不变性模型几乎不会被满足(Marsh et al., 2018)
- **事后模型修正：**
 - 修正过程繁琐
 - 参数的释放会受到研究者主观性的影响，增加了一类错误率，降低了结论的可重复性
 - 基于修正指数的方法难以找到全局最优的模型(Chou & Bentler, 1990)
 - 容易导致对跨组因子均值差异的有偏估计(Marsh et al., 2018)

贝叶斯渐近测量不变性

- **贝叶斯方法 vs 频率学方法**

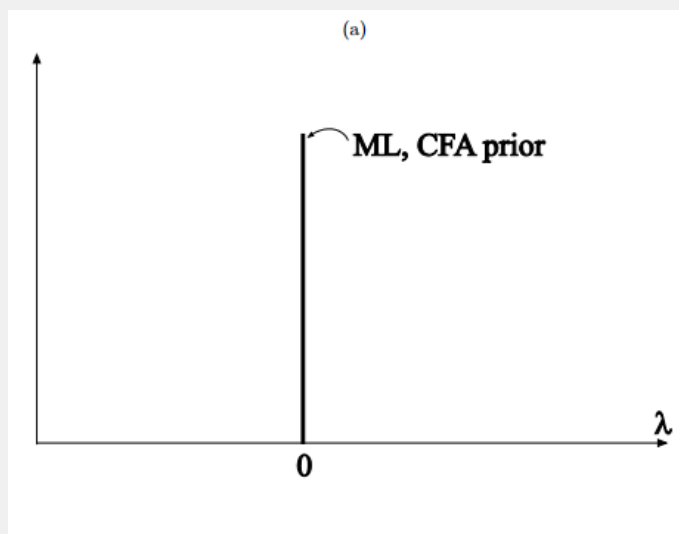
- 频率学方法将未知参数视为常数，根据样本参数估计总体参数
- 贝叶斯方法则将其视为随机变量，结合样本数据和先验信息得到未知参数的后验分布(王孟成等, 2017)

- **先验信息**

- 灵活地纳入先验信息（如，前人研究或预实验结果）
- 有效的先验信息可以带来更准确的参数估计(Yuan & MacKinnon, 2009)

先验信息

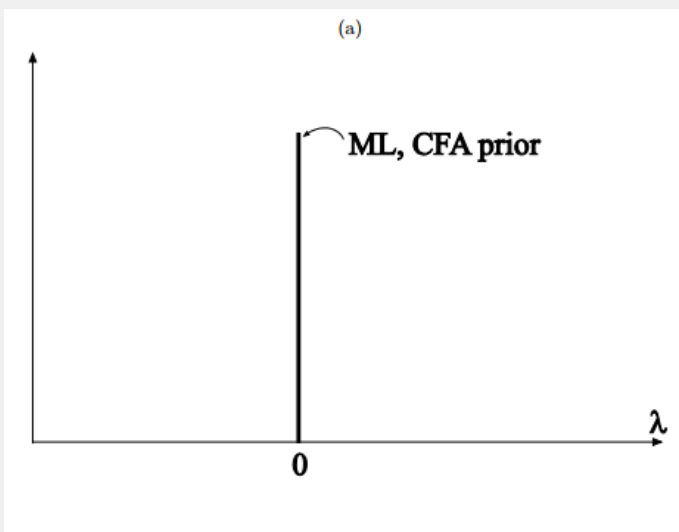
- 例如，结构方程模型中被限制为0的参数就可以被视为提供了均值为0，方差为0的先验分布。



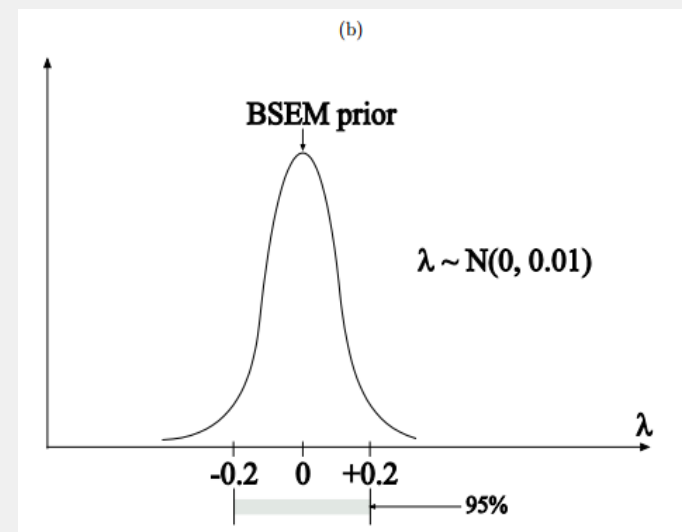
贝叶斯渐近测量不变性

- Muthén和Asparouhov(2013)

- 对载荷和截距的跨组差异提供均值为0、方差极小的先验分布
- 放宽传统方法对其的严格限制，允许载荷和截距存在微小的跨组差异（渐近不变）
- 模型拟合良好的同时实现对结构模型参数的准确估计



----- (10) -----



贝叶斯渐近测量不变性

优势

- 避免了传统方法限制过于严格而带来的一系列问题：
- 避免了修正时重复利用同一批数据可能导致的更高的一类错误率、有偏的参数估计等问题

局限

应用步骤、报告标准尚未规范

- 模型拟合标准
- 先验方差的选择(De Bondt & Van Petegem, 2015; Cieciuch et al., 2017)
 - 不同方差的先验分布对模型参数估计的影响未知
 - “微小差异”的定义取决于数据条件

研究目的

- 01 模拟研究：探究在不同数据条件下，不同先验方差会对因子均值差异的估计产生怎样的影响
- 02 基于模拟研究结果给应用研究者提供有效的建议
- 03 实证研究：验证模拟研究结论的有效性
- 04 与传统的多组CFA方法进行深入对比，展示渐近测量不变性方法的优势



中山大學心理學系

SUN YAT-SEN UNIVERSITY DEPARTMENT OF PSYCHOLOGY

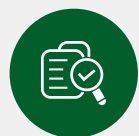
模拟研究

PART TWO

在不同数据条件下，不同先验方差的影响

数据生成

采用R软件中mass、stats和truncnorm软件包生成以下144个条件下各100个数据集用于模型估计



模型大小

(Asparouhov & Muthén, 2014)

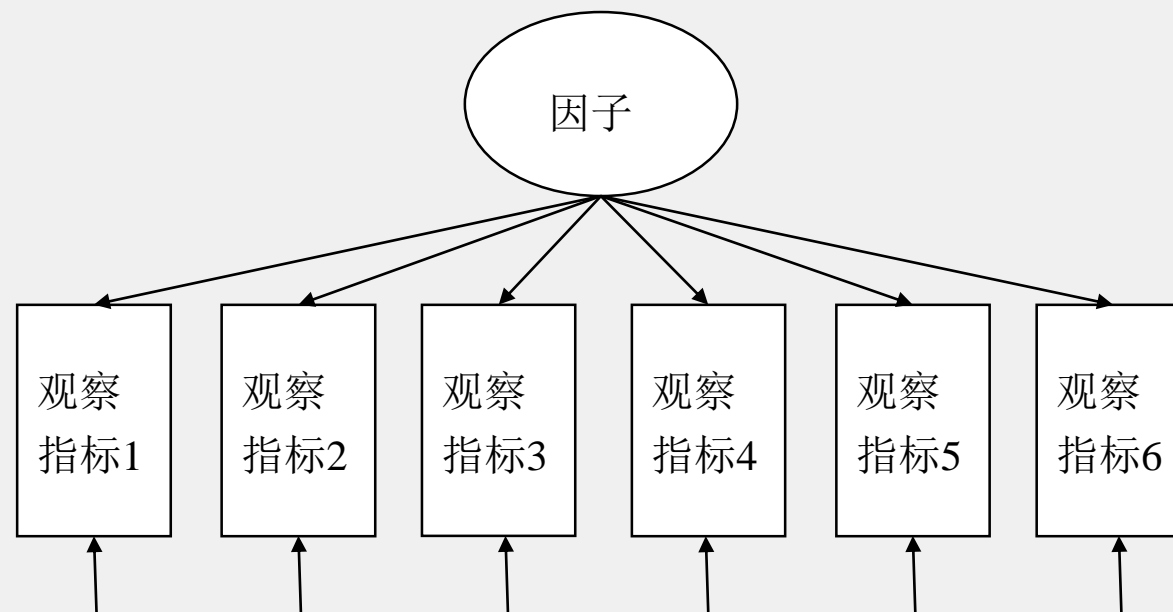
- 观察指标数= 6/10



样本-条目比

(李锡钦, 2011)

- $\frac{\text{每组样本量}}{\text{观察指标数}} = 10/20/40/60$



数据生成



组数

(Asparouhov & Muthén, 2014)

- Num of Groups = 2/3/15
- 第1组（参照组）的因子采用 $N(0,1)$
- 第2组 $N(1, 1)$
- 第3组 $N(-1, 1)$
- 15组条件：
 - 第1（参照组）、4、7、10、13组因子 $N(0, 1)$ 分布
 - 第2、5、8、11、14组 $N(1, 1)$
 - 第3、6、9、12、15组 $N(-1, 1)$



参数跨组差异程度

(Kim et al., 2017)

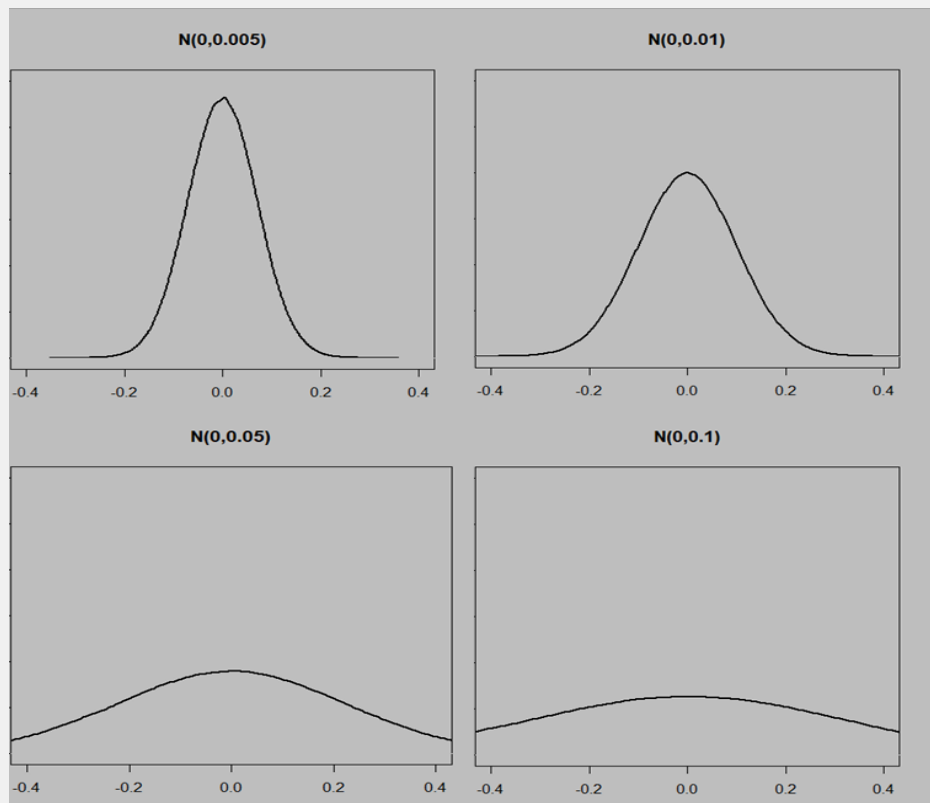
- 载荷 & 截距
- $N(0.7, 0.0001)$ & $N(0, 0.0001)$
- $N(0.7, 0.0025)$ & $N(0, 0.0025)$
- $N(0.7, 0.005)$ & $N(0, 0.005)$
- $N(0.7, 0.01)$ & $N(0, 0.01)$
- $N(0.7, 0.05)$ & $N(0, 0.05)$
- $N(0.7, 0.1)$ & $N(0, 0.1)$

模型估计

采用R软件中MplusAutomation软件包和Mplus8.0软件

贝叶斯渐近测量不变性

- 针载荷和截距的跨组差异提供了 $N(0,0.005)$, $N(0,0.01)$, $N(0,0.05)$, $N(0,0.1)$ 四种先验分布
- 分别允许载荷和截距参数的跨组差异值有95%的概率落在 $(-0.14, 0.14)$, $(-0.2, 0.2)$, $(-0.45, 0.45)$, $(-0.63, 0.63)$ 之中



模拟研究结果

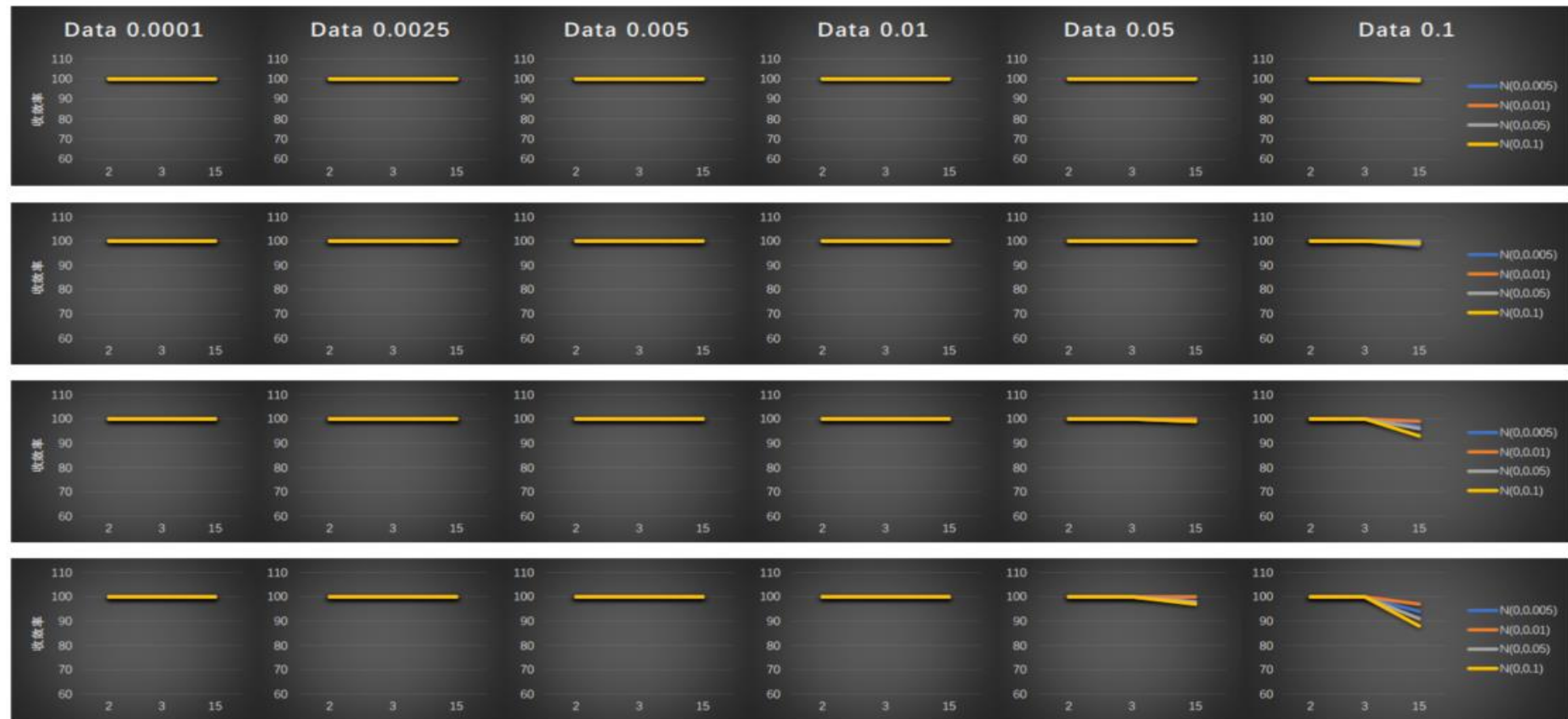
步骤

- 1 收敛率：在每个实验条件中重复 100次分析下，完成数据分析的比率。
- 2 模型拟合：每个实验条件下进行100次分析后的 PP_p 值均值是否大于0.1
- 3 参数估计（因子均值差异）

模型收敛率

六指标模型-收敛率

样本条目比=10

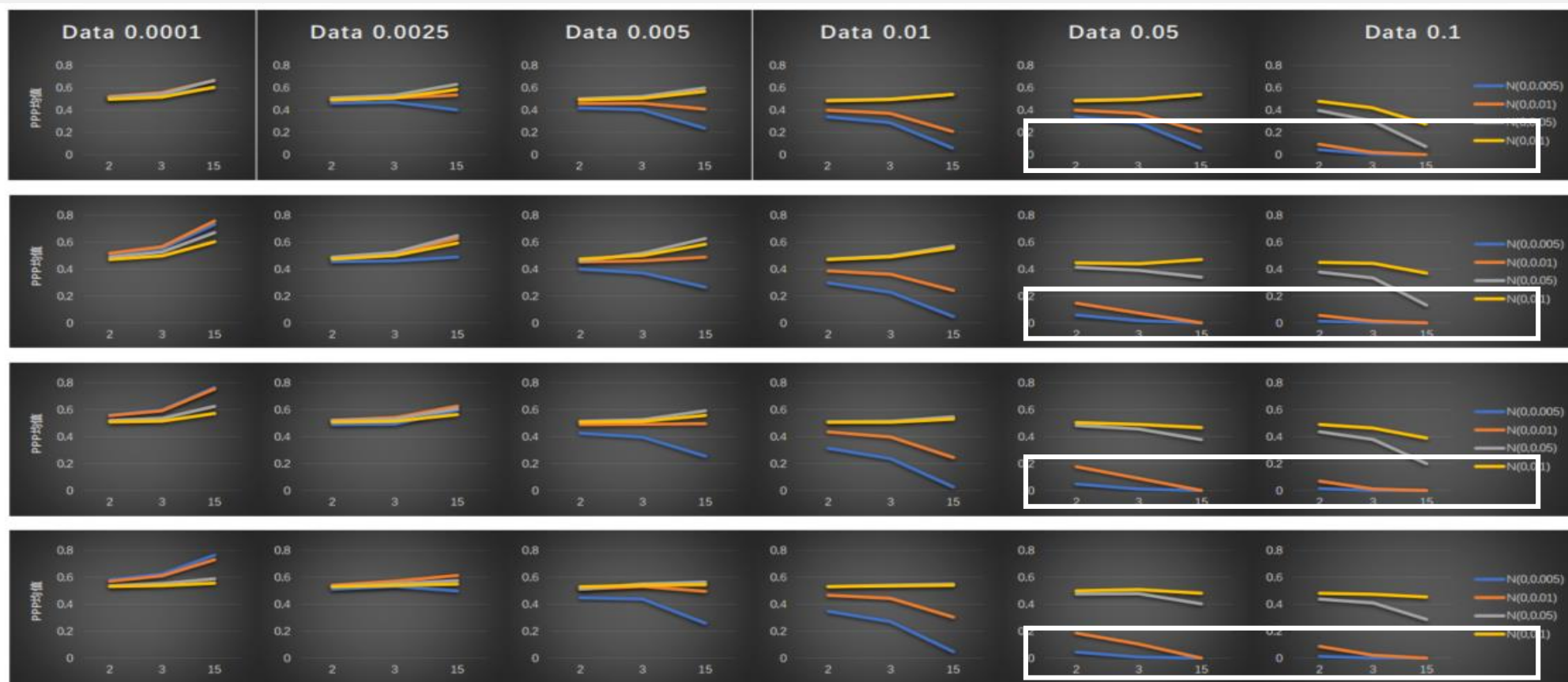


样本条目比=60

模型拟合情况-PP_p均值

六指标模型-模型拟合情况

样本条目比 1:10



样本条目比 1:40

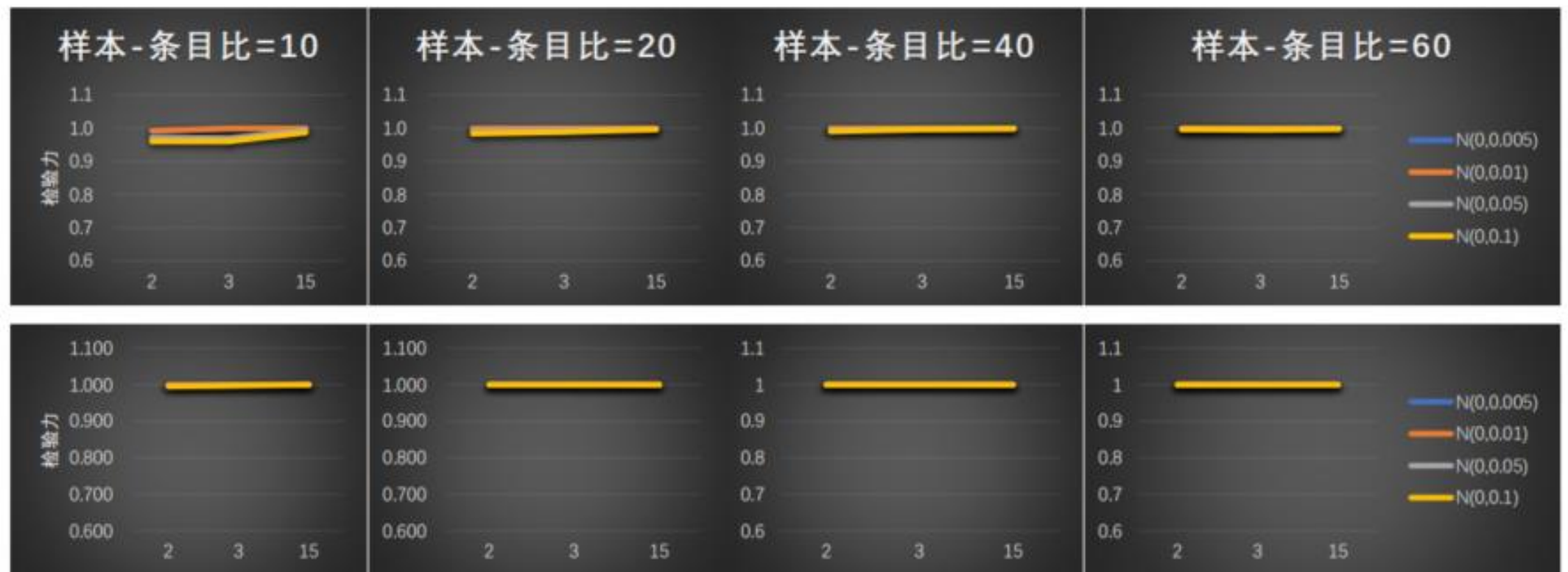
评估因子均值差异估计准确性的指标

(Cham, West, Ma, & Aiken, 2012; Marsh, Wen, & Hau, 2004; Cheung & Lau, 2017)

	评估指标名称	标准
1	检验力(Power)	$\geq 80\%$
2	标准误估计准确性(Standard error ratio)	$0.9 \leq SE\ ratio \leq 1.1$
3	覆盖率(Coverage rate)	$\geq 90\%$
4	系数估计相对偏差(Relative Bias)	$\leq 10\%$
5	一类错误率(Type I error)	$\leq 5\%$

检验力

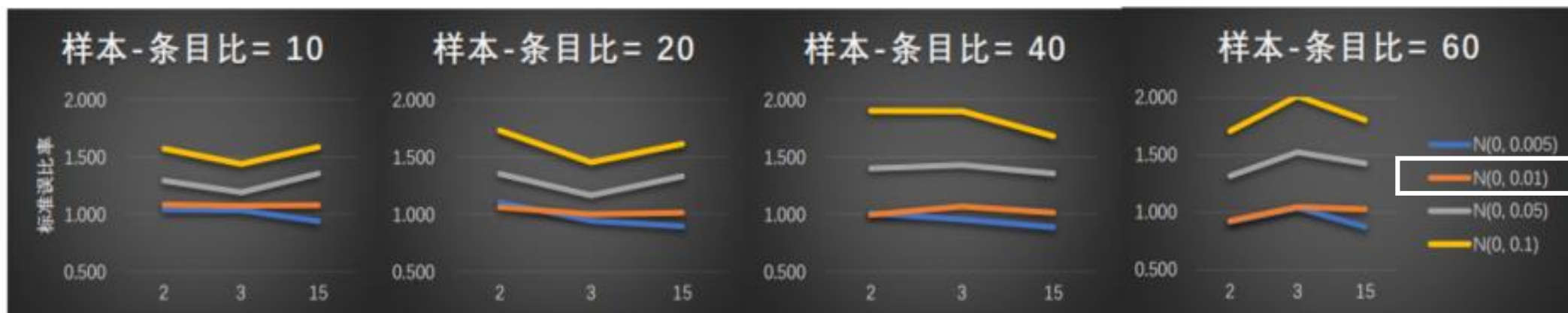
不同先验分布下因子均值差异估计的检验力
($\geq 80\%$)



标准误比率

不同先验分布下因子均值差异估计的标准误比率
($0.9 \leq SE \text{ ratio} \leq 1.1$)

观察指标数=9

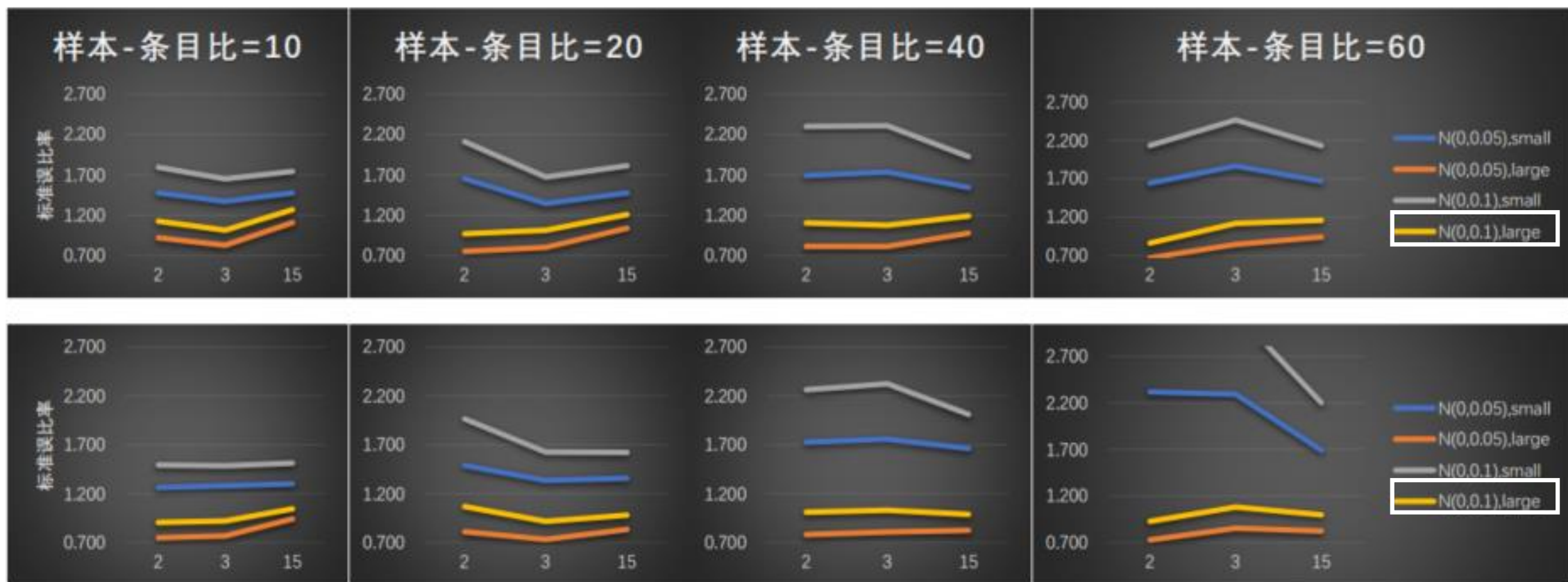


观察指标数=10



标准误比率

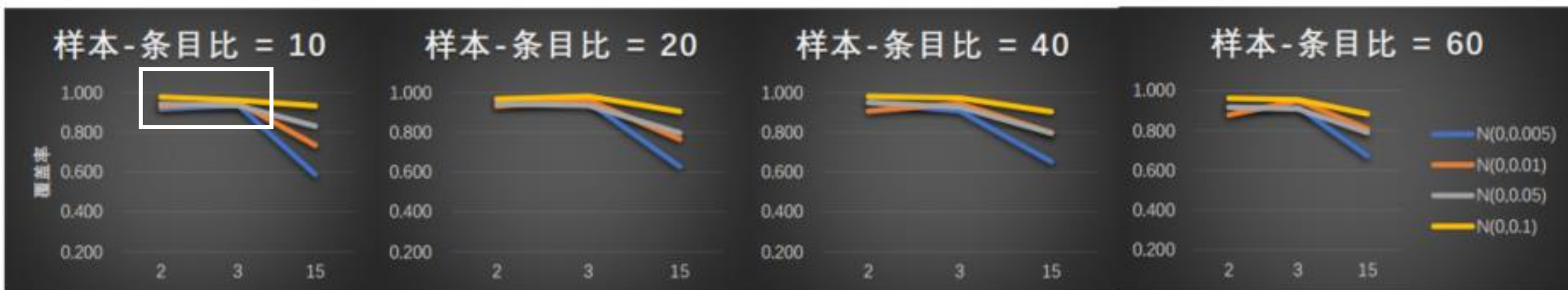
$N(0,0.05)$ 和 $N(0,0.1)$ 的先验分布在不同跨组差异幅度下的表现



覆盖率

不同先验分布下因子均值差异估计的覆盖率
($\geq 90\%$)

观察指标数 9



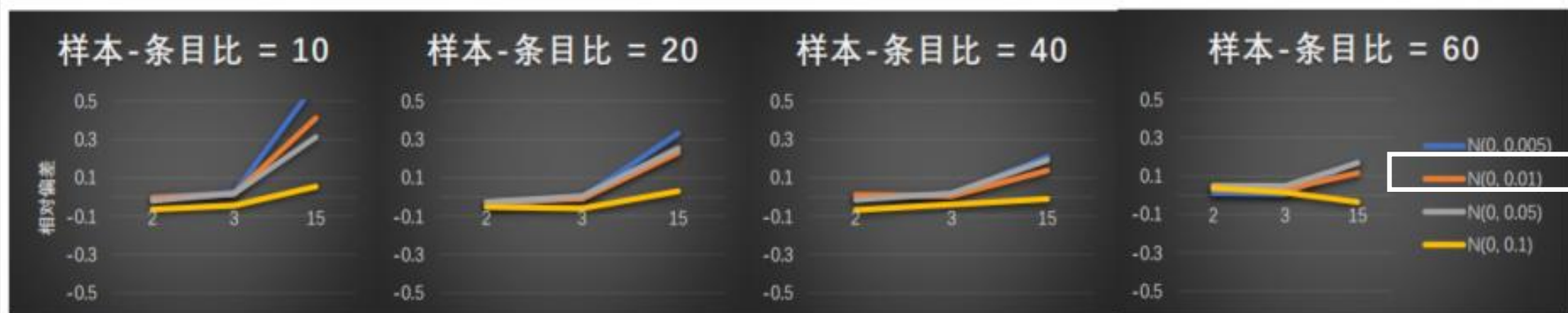
观察指标数 10



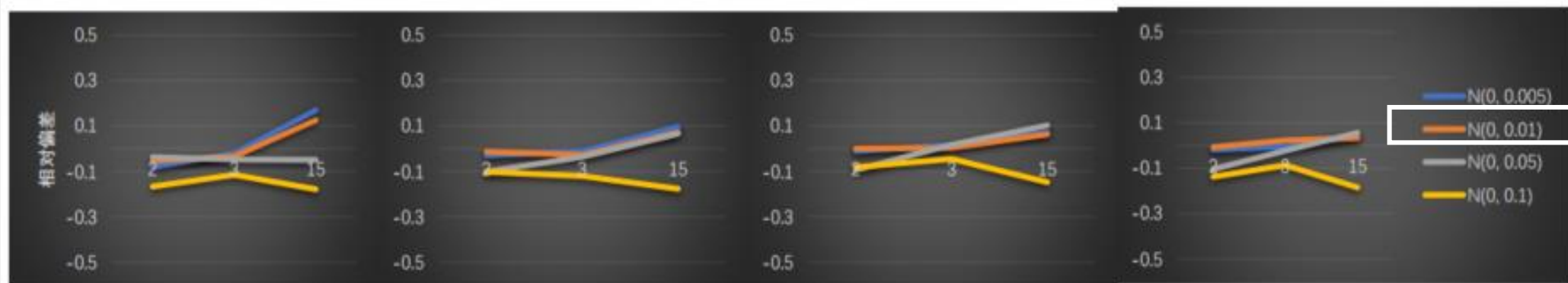
系数估计相对偏差

不同先验分布下因子均值差异估计的相对偏差
($\leq 10\%$)

观察指标数 9



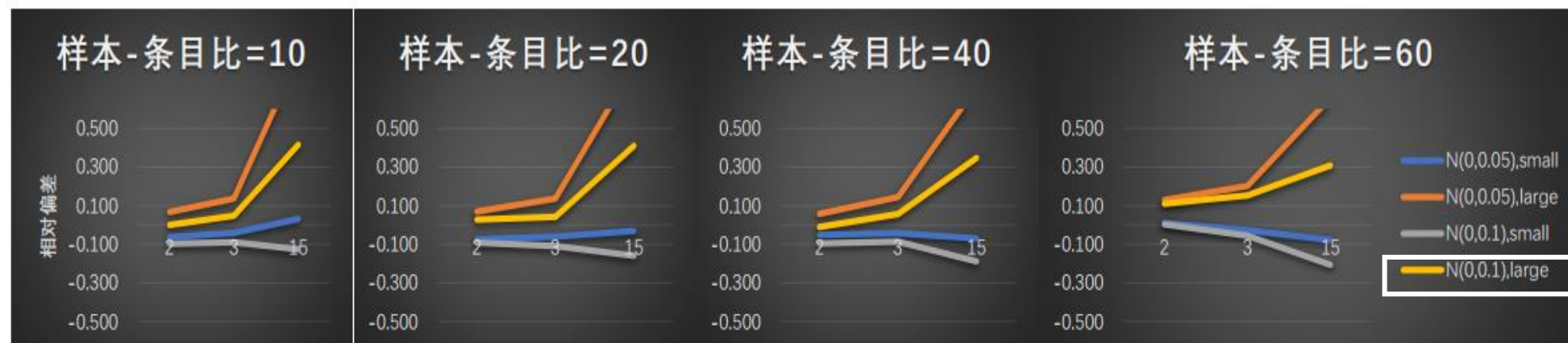
观察指标数 10



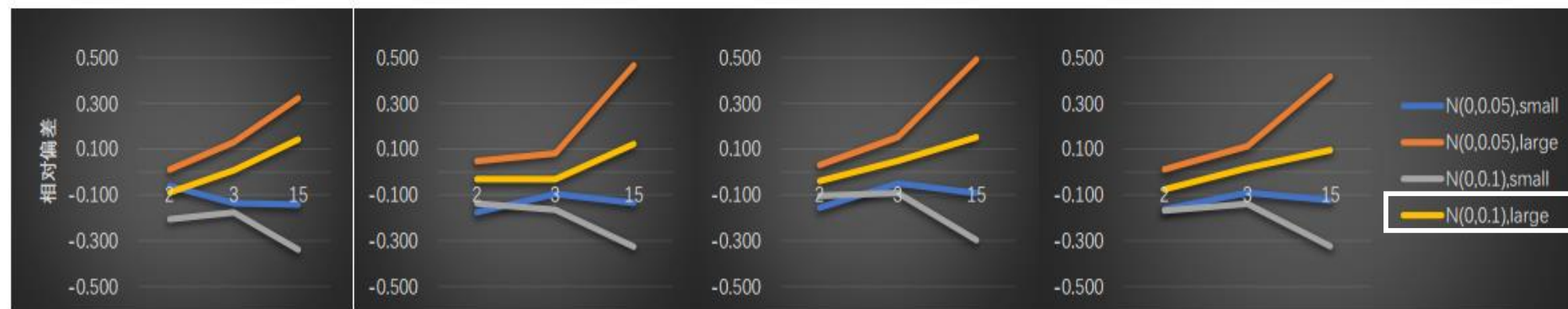
系数估计相对偏差

$N(0,0.05)$ 和 $N(0,0.1)$ 的先验分布在不同跨组差异幅度下的表现

观察指标数
9



观察指标数
10

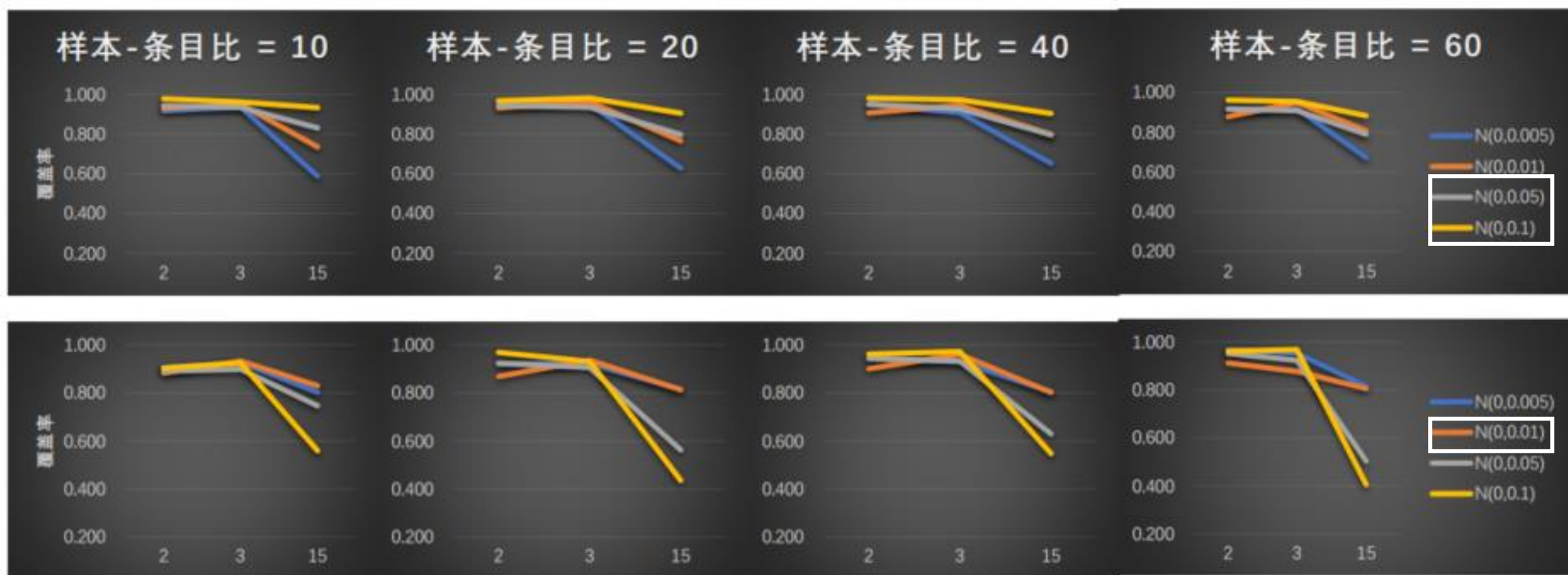


结论1

- 组数少，跨组差异幅度小：
 - $N(0,0.01)$ 可以获得高检验力、高覆盖率、准确的标准误和参数估计结果
- 组数少，跨组差异幅度大：
 - $N(0,0.1)$ 可以获得高检验力、高覆盖率、准确的标准误和参数估计结果
- 建议在组数较少（2组/3组）的条件下
 - ✓ 首先采用 $N(0,0.01)$ 先验分布进行建模
 - ✓ 如果模型拟合较差则采用 $N(0,0.1)$

覆盖率

不同先验分布下因子均值差异估计的覆盖率



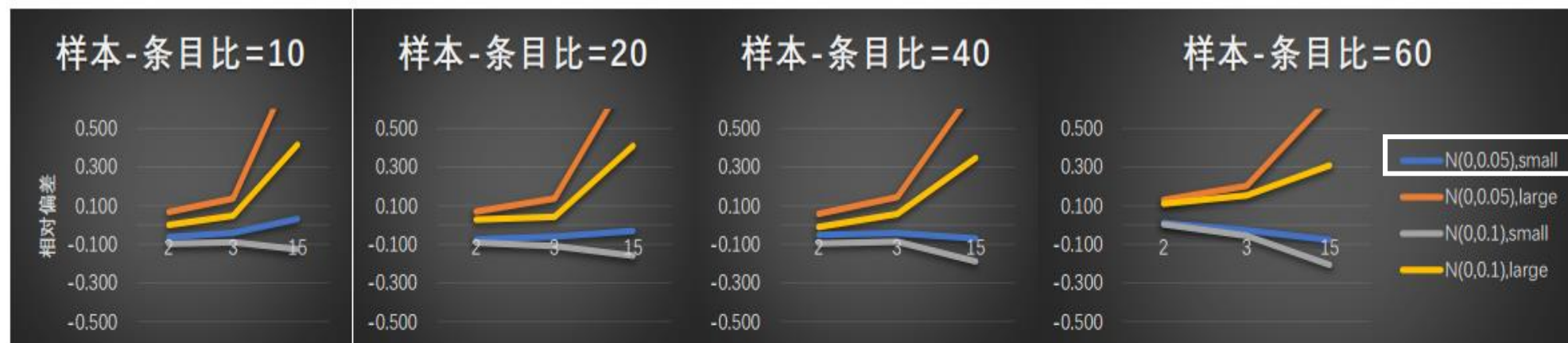
观察指标数 9

观察指标数 10

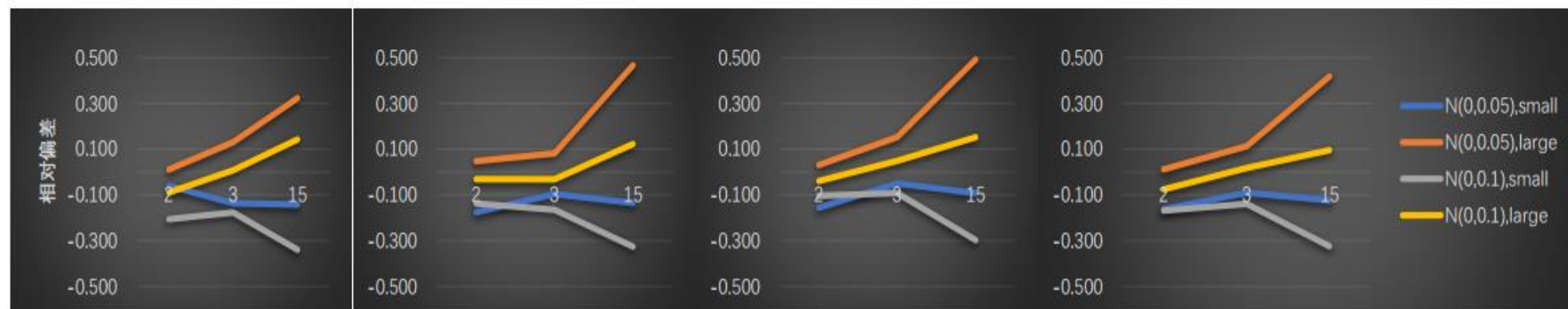
系数估计相对偏差

$N(0,0.05)$ 和 $N(0,0.1)$ 的先验分布在不同跨组差异幅度下的表现

观察指标数 9

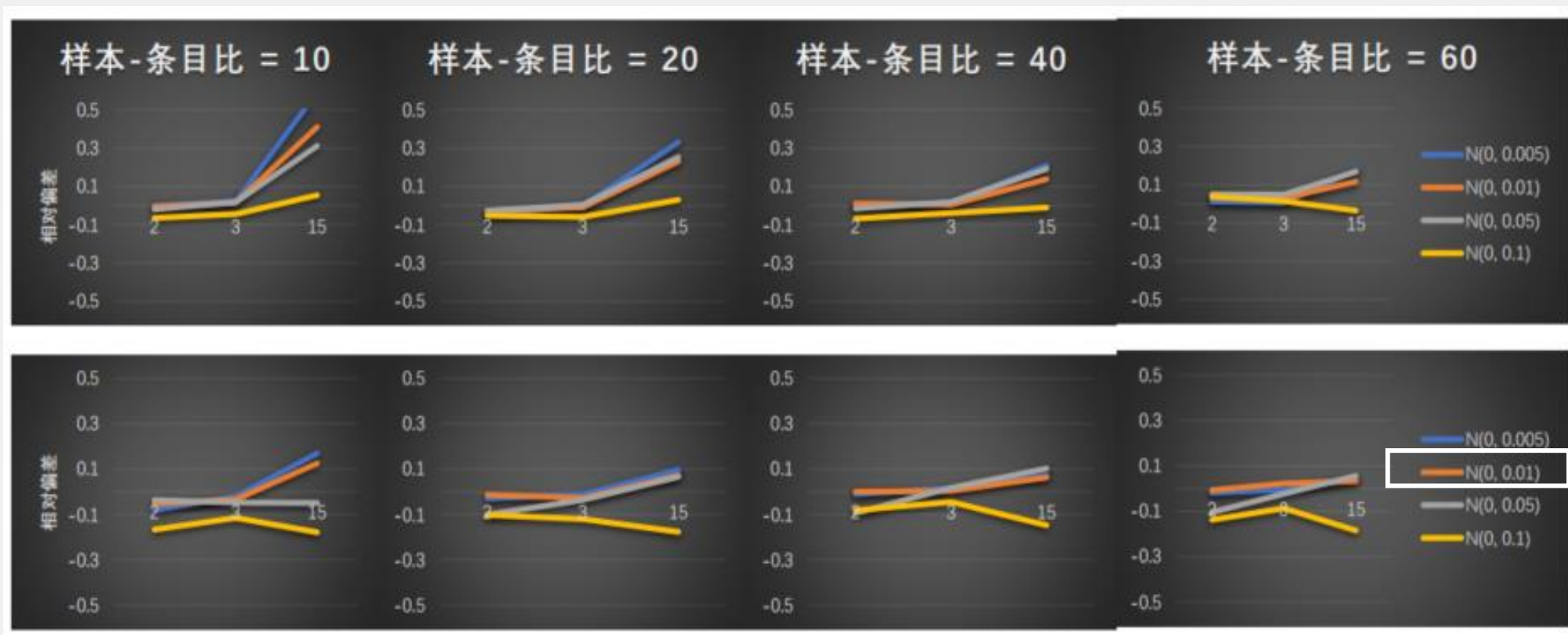


观察指标数 10



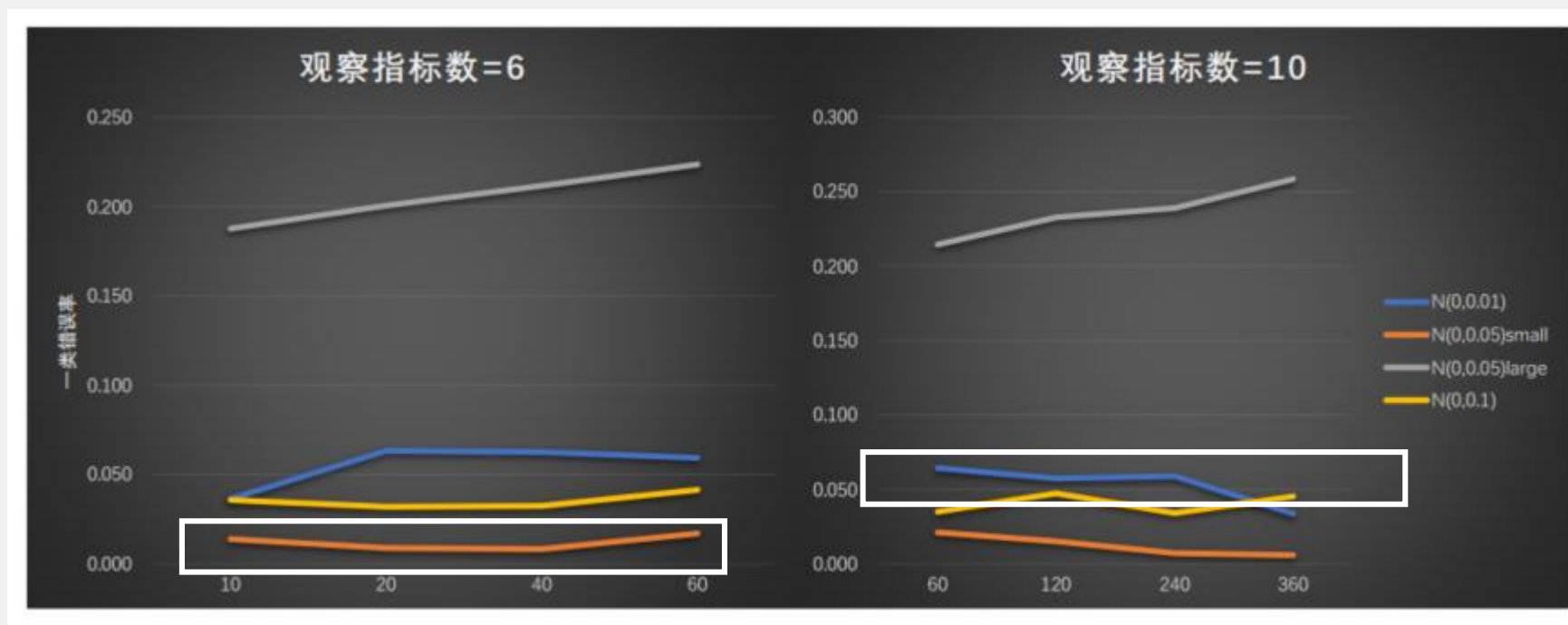
系数估计相对偏差

不同先验分布下因子均值差异估计的相对偏差



一类错误率

不同先验分布下因子均值差异估计的一类错误率(15组条件下)
($\leq 5\%$)



研究结论与建议

数据条件	建议
组数 ≤ 3	首先采用 $N(0,0.01)$ 先验分布进行建模 如果模型拟合较差则采用 $N(0,0.1)$ 分布
组数 > 3 指标数 < 10	首先采用 $N(0,0.01)$ 先验分布进行建模 如果模型拟合良好，则采用 $N(0,0.05)$ 的先验分布进行建模
组数 > 3 指标数 ≥ 10 样本-条目比 > 10	采用 $N(0,0.01)$ 先验分布进行建模



中山大學心理學系
SUN YAT-SEN UNIVERSITY DEPARTMENT OF PSYCHOLOGY

实证研究

PART THREE

不同性别在社会支持量表三个维度中的表现差异

贝叶斯渐近测量不变性 vs 多组验证性因子分析

研究方法 – 被试与工具

- 测量工具

- 大学生社会支持评定量表（叶悦妹，戴晓阳，2008）：17个条目，5点计分
- 包括主观支持、客观支持和支持利用度三个维度，每个维度分别有5题、6题和6题

- 被试

- 共收集353名本科大学生被试。其中男生89名
- 平均年龄为19.55岁($SD = 1.45$, range:17-24)

- 分析工具：SPSS 20.0 和 *Mplus* 8.0

研究方法 – 数据分析流程与结果

信度分析

多组CFA

贝叶斯渐近测量不变性

- 内部一致性系数
 - 问卷内部一致性系数为0.928
 - 三个子维度的内部一致性系数分别为0.892、0.818 和 0.881
 - 信度良好
- 组合信度
 - 三个子维度的组合信度分别为0.894、0.819 和 0.882
 - 信度良好



研究方法 – 数据分析流程与结果

信度分析

多组CFA

贝叶斯渐近测
量不变性

- 峰度偏度分析
 - 各条目的偏度在-1.107到-1.43之间，峰度在-0.705到1.534之间
 - 数据中不含缺失值
 - 采用极大似然法(Maximum Likelihood, ML)进行模型估计
- 模型拟合比较
 - $|\Delta CFI| < 0.01$ (Little & Card, 2013)

研究方法 – 数据分析流程与结果

多组验证性因子分析模型拟合及模型比较

信度分析

多组CFA

贝叶斯渐近测
量不变性

自由估计2个载荷参数

模型	χ^2	df	CFI	TLI	SRMR	RMSEA	90%CI of RMSEA	Δ CFI	Pass?
形态不变性	494.817	230	0.921	0.907	0.064	0.081	[0.071,0.091]		
载荷不变性	521.450	244	0.917	0.908	0.075	0.080	[0.071,0.090]	-0.04	No
部分载荷不变性	512.170	242	0.920	0.910	0.072	0.080	[0.070,0.089]	-0.01	Yes
截距不变性	560.015	256	0.909	0.904	0.083	0.082	[0.073,0.091]	-0.11	No
部分截距不变性	526.292	253	0.919	0.913	0.076	0.078	[0.069,0.088]	-0.01	Yes

自由估计3个截距参数

研究方法 – 数据分析流程与结果

信度分析

多组CFA

贝叶斯渐近测
量不变性

分析流程

评价标准

1 形态不变性模型

PP_p在0.5左右, 后验预测区间不包括0

2 载荷和截距不变性模型
Prior: $N(0, 0.01)$

PP_p在0.5左右, 后验预测区间不包括0

3 敏感性分析 (Kim et al., 2017)
Prior: $N(0, 0.05)$

与模型2的拟合没有明显差异: $|\Delta DIC| < 7$
(Spiegelhalter, Best, Carlin, & Van Der Linde, 2002)

研究方法 – 数据分析流程与结果

信度分析

多组CFA

贝叶斯渐近测量不变性

模型	PPp	95%C.I.	BIC	DIC	ΔDIC
形态不变性	0.624	[-86.681, 62.040]	14723.613	13131.384	
载荷和截距不变性 (先验方差0.01)	0.664	[-90.916, 60.739]	14766.800	13113.782	
载荷和截距不变性 (先验方差0.05)	0.674	[-90.851, 57.596]	14757.127	13111.530	-2.252
部分测量不变性 (Muthén & Asparouhov, 2013; van de Schoot et al., 2013)	0.681	[-93.554, 58.890]	14764.332	13112.122	

方法比较

信度分析

多组CFA

贝叶斯渐近测
量不变性

与Muthén和Asparouhov(2012)及van de Schoot等人(2013)的研究结果一致的是：渐近测量不变性方法对因子均值差异的估计结果与传统方法相差不大。

多组CFA	渐近测量不变性
$Diff = 0.389, p = 0.002$	$Diff = 0.394, p = 0.004$
$Diff = 0.566, p < 0.001$	$Diff = 0.443, p < 0.001$
$Diff = 0.410, p = 0.003$	$Diff = 0.559, p < 0.001$

方法比较

传统方法

- 对模型限制过于严格
- 模型比较和修正过程都更为繁琐
- 发现更多的违背不变性的载荷和截距参数

渐近测量不变性

- 模拟研究所提供的先验分布建议可以有效地放宽传统方法对模型的严格限制
- 操作简单，拟合良好
- 敏感性分析帮助研究者得到更可靠的结果



中山大學心理學系
SUN YAT-SEN UNIVERSITY DEPARTMENT OF PSYCHOLOGY

总 讨 论

PART FOUR

总结与展望

理论意义与应用价值

01 模拟研究

- 为应用研究者提供了准确有效的建议
- 自贝叶斯结构方程建模放宽模型限制的思想被提出以来(Muthén & Asparouhov, 2012), 一直以来对于小方差先验的选取都没有规范的标准, 希望本研究能够促进这种建模思想的完善

02 实证研究: 对比传统方法与渐近测量不变性方法

- 相比于传统方法, 贝叶斯方法在模型拟合上具有极大的优势
- 不需要进行多次事后模型修正

研究不足以及未来研究方向

研究局限

- 难以覆盖实际数据中的全部情况
- 如，组数更多的数据、样本量跨组不等的情况...

未来研究方向

- 模型拟合评估标准
- 发展检测实际数据的参数跨组差异幅度的研究工具—敏感性分析
- 部分渐近测量不变性：渐近测量不变性方法作为一种检测违背不变性的参数的工具时的表现



— 感谢 —
Thanks

特别感谢

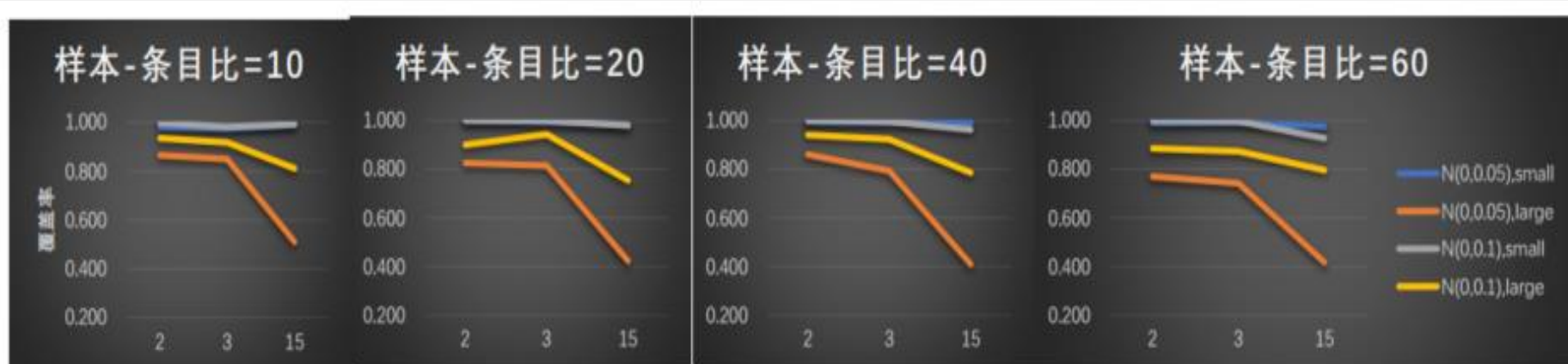
潘俊豪老师, Pan Lab

本文摘要已被2019年国际心理测量研讨会(International Meeting of Psychometric Society)接收为分组口头报告

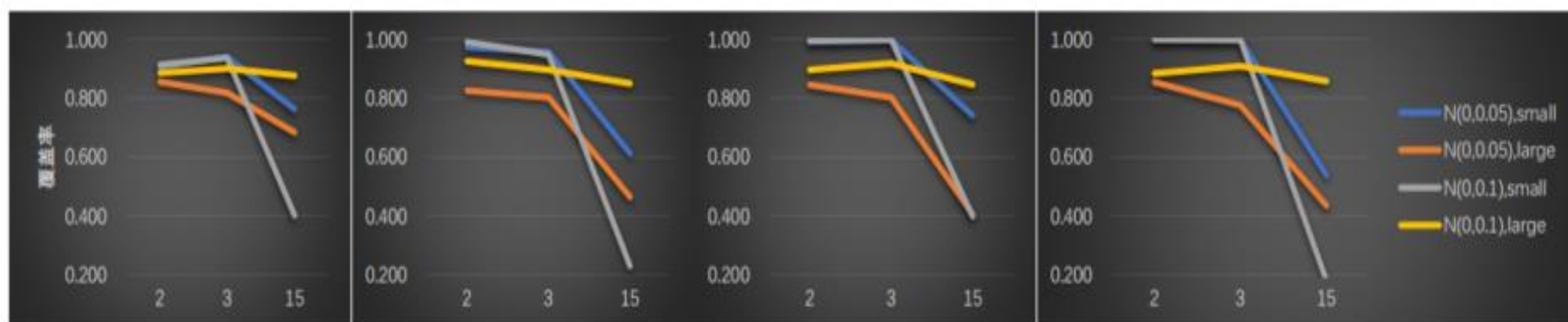
覆盖率

$N(0,0.05)$ 和 $N(0,0.1)$ 的先验分布在不同跨组差异幅度下的表现

观察指标数 9



观察指标数 10



覆盖率

十指标模型

样本-条目比		10			20			40			60		
先验	数据	G=2	G=3	G=15	G=2	G=3	G=15	G=2	G=3	G=15	G=2	G=3	G=15
N(0,0.005)	0.005	0.900	0.925	0.639	0.930	0.925	0.629	0.960	0.935	0.559	0.970	0.960	0.605
	0.01	0.870	0.905		0.880	0.870		0.870	0.845		0.940	0.875	
	0.05												
	0.1												
N(0,0.01)	0.005	0.920	0.930	0.854	0.930	0.925	0.880	0.970	0.960	0.881	0.980	0.980	0.889
	0.01	0.870	0.920	0.579	0.910	0.915	0.524	0.910	0.915	0.423	0.970	0.925	0.476
	0.05	0.770			0.620			0.660			0.630	0.490	
	0.1												

样本-条目比		10			20			40			60		
先验	数据	G=2	G=3	G=15	G=2	G=3	G=15	G=2	G=3	G=15	G=2	G=3	G=15
N(0,0.005)	0.005	0.422	0.378	0.242	0.461	0.402	0.252	0.44	0.371	0.198	0.455	0.42	0.253
	0.01	0.323	0.228	0.044	0.342	0.214	0.022	0.323	0.223	0.021	0.355	0.273	0.033
	0.05	0.032	0.002	0	0.023	0.002	0	0.021	0.002	0	0.033	0.005	0
	0.1	0.002	0	0	0.002	0	0	0.004	0	0	0.002	0	0
N(0,0.01)	0.005	0.478	0.467	0.443	0.526	0.512	0.528	0.509	0.472	0.454	0.519	0.512	0.509
	0.01	0.42	0.368	0.217	0.464	0.397	0.255	0.461	0.4	0.242	0.487	0.456	0.327
	0.05	0.117	0.038	0	0.139	0.049	0	0.176	0.074	0	0.211	0.104	0.001
	0.1	0.009	0.002	0	0.032	0.002	0	0.061	0.008	0	0.08	0.015	0

- PP_p 均值 ≥ 0.1 的标准较为武断
- 贝叶斯渐近测量不变性的模型拟合标准有待深入探究



参考文献

- 侯杰泰, 温忠麟, 成子娟. (2004). *结构方程模型及其应用*. 北京: 教育科学出版社.
- 李锡钦. (2011). *结构方程模型: 贝叶斯方法* (蔡敬衡, 潘俊豪, 周影辉译). 北京: 高等教育出版社.
- 欧阳湘子, 田伟, 辛涛, & 詹沛达. (2016). IRT框架下追踪数据的测量不变性分析——以4至5岁儿童认知能力测验为例. *心理科学*(3), 606-613.
- 王孟成, 邓倩文, 毕向阳. (2017). 潜变量建模的贝叶斯方法. *心理科学进展*, 25(10), 1682-1695.
- 王济川, 王小倩, 姜宝法. (2011). *结构方程模型: 方法与应用*. 北京: 高等教育出版社.
- 徐万里. (2008). 结构方程模式在信度检验中的应用. *统计与信息论坛*, 23(7), 9-13.
- 叶悦妹, 戴晓阳. (2008). 大学生社会支持评定量表的编制. *中国临床心理学杂志*, 16(5): 465-468.
- 张沥今, 陆嘉琦, 魏夏琰, 潘俊豪. (2018). 贝叶斯结构方程模型及其研究现状. *心理科学进展*, doi:10.12074/201812.00865.
- Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(4), 495-508.
- Binder, K., Heermann, D., Roelofs, L., Mallinckrodt, A. J., & McKay, S. (1993). Monte carlo simulation in statistical physics. *Computers in Physics*, 7(2), 156.
- Billiet, J. (2013). Quantitative methods with survey data in comparative research. *Urology*, 76(76), S2-S3.
- Chou, C. P., & Bentler, P. M. (1990). Model modification in covariance structure modeling: A comparison among likelihood ratio, Lagrange multiplier, and Wald tests. *Multivariate Behavioral Research*, 25(1), 115-136.
- Cieciuch, J., Davidov, E., Algesheimer, René., & Schmidt, P. (2017). Testing for approximate measurement invariance of human values in the European social survey. *Sociological Methods & Research*, 47(4), 665-686.
- Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 45-97.

参考文献

- De Bondt, N., & Van Petegem, P. (2015). Psychometric Evaluation of the Overexcitability Questionnaire-Two Applying Bayesian Structural Equation Modeling (BSEM) and Multiple-Group BSEM-Based Alignment with Approximate Measurement Invariance. *Frontiers in Psychology*, 6, 1963.
- Fong, T., & Ho, R. (2014). Testing gender invariance of the hospital anxiety and depression scale using the classical approach and bayesian approach. *Quality of Life Research*, 23(5), 1421-1426.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian data analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall.
- Hallquist, M. N. & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling*, 1-18.
- Hoogland, J. J. , & Boomsma, A. . (1998). Robustness studies in covariance structure modeling: an overview and a meta-analysis. *Sociological Methods & Research*, 26(3), 329-367.
- Jak, S., Oort, F. J., & Dolan, C. V. (2013). A test for cluster bias: detecting violations of measurement invariance across clusters in multilevel data. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(2), 265-282.
- Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement invariance testing with many groups: a comparison of five approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(4), 524-544.
- Lee, S. Y, & Song, X. Y. (2012). *Basic and advanced Bayesian structural equation modeling*. John Wiley & Sons, Ltd.
- Little, T. D., & Card, N. A. (2013). *Longitudinal structural equation modeling*. The Guilford Press.
- Lu, Z. H., Chow, S. M., & Loken, E. (2016). Bayesian factor analysis as a variable-selection problem: alternative priors and consequences. *Multivariate Behavioral Research*, 51(4), 519–539.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalizing on chance. *Psychological Bulletin*, 111, 490–504. doi:10.1037/0033-2909.111.3.490



参考文献

- Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B. , Asparouhov, T. , & Muthén B, et al. (2018). What to do when scalar invariance fails: the extended alignment method for multi-group factor analysis comparison of latent means across many groups. *Psychological Methods*, 23(3), 524-545.
- Meade, A. W., & Lautenschlager, G. J. (2004). A Monte-Carlo study of confirmatory factor analytic tests of measurement equivalence/invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 11(1), 60-72.
- Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313–335.
- Muthén, L. K., & Muthén, B. O. (1998-2019). *Mplus user's guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(4), 599-620.
- Pan, J., Ip, E. H., & Dubé, L. (2017). An alternative to post hoc model modification in confirmatory factor analysis: the Bayesian lasso. *Psychological Methods*, 22(4), 687–704.
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. R foundation for statistical computing. Retrieved from <http://www.Rproject.org>.
- Raykov, T. (2001). Bias of coefficient alpha fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25(1), 69-76.
- Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational & Psychological Measurement*, 74(1), 31-57.
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54, 371–384.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583-639.



参考文献

- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: review of practice and implications. *Human Resource Management Review*, 18(4), 210–222.
- 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.
- van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., & Muthén, B. (2013). Facing off with scylla and charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology*, 4, 770.
- van de Schoot, R., Winter, S. D., Ryan, O., Zondervanzwijnenburg, M., & Depaoli, S. (2017). A systematic review of Bayesian articles in psychology: the last 25 years. *Psychological Methods*, 22(2), 217-239.
- West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications*. (pp. 56-75). Thousand Oaks, CA: Sage Publications, Inc.
- Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, 14(4), 301-322.