

# 验证性因素分析中模型拟合的判断<sup>\*</sup>

郭庆科 王炜丽 陈雪霞 韩 丹

(辽宁师范大学 心理学系, 大连 116029)

**摘 要** 采用与 Hu & Bentler 不同的设计进行模拟研究, 发现在他们推荐的 8 个指数中, NNFI、CFI、IFI 在所有样本量条件下几乎都能得到最低的  $\alpha$  与  $\beta$  错误率, 而且最佳界值都在 0.95 左右, 属于优良的指数。但 Mc 与 Gamma Hat 的表现却不好, 因此不建议使用。他们建议的 2 指标策略被证明能降低  $\alpha$  与  $\beta$  错误率, 在样本量较小时尤其值得尝试。

**关键词** 结构方程模型 模型拟合 拟合指数  $\alpha$  与  $\beta$  错误

**中图分类号** B841.2

**文献标识码** A

**文章编号** 1003-5184(2007)04-0083-05

## 1 问题的提出

### 1.1 拟合指数及其意义

心理与教育研究中结构方程模型(SEM)常被用来检验研究者提出的理论模型是否得到了数据的支持, 其判断标准是 LISREL 等统计软件输出的拟合指数<sup>[1-3]</sup>。但是模型拟合并没有统一的标准, 迄今公布的拟合指数都存在不同程度的问题, 根据不同指数常常会得出很不相同的结论。选择哪些指数作为判断模型拟合的标准, 各拟合指数的临界值在多大时才能避免拒绝正确模型和接受错误模型等问题还没有得到很好的回答。

### 1.2 关于拟合指数的模拟研究

只有在判断模型拟合时选择了好的拟合指数, 才能避免错误拒绝正确的模型(犯  $\alpha$  错误或 I 型错误)或错误接受不正确的模型(犯  $\beta$  错误或 II 型错误)。由于在实践中真模型是不知道的, 要检验拟合指数的性能只能通过模拟研究<sup>[4, 5]</sup>。历史上有几个模拟研究(如 Bollen, Bentler, McDonald & Marsh, Marsh & Hau)非常有影响, 其结论对拟合指数的选用产生了重要影响<sup>[3, 6-9]</sup>。

Hu & Bentler 在 1998 和 1999 年所做的研究当属本领域最有影响的。他们系统探讨了各拟合指数在不同估计方法、不同分布等条件下的表现, 结果发现 1)以极大似然估计为基础的拟合指数优于以 GLS 和 ADF 为基础的指数; 2)SRMR、NNFI、BL89(或 IFI)、RNI、CFI、Gamma Hat、Mc、RMSEA 性能较好, 值得推荐, 其中 SRMR 对错误设定的因子相关敏感, 而其他 7 个指数则对错误设定的因子载荷敏感; 3)在 NNFI、IFI、RNI、CFI 和 Gamma Hat  $\approx 0.95$ , Mc  $\approx 0.9$ ,

SRMR  $\approx 0.08$ , RMSEA  $\approx 0.06$  时两类错误率都不高, 是值得推荐的界值标准。但结论 3)在很多条件下仍然容易导致错误的结论, 因此他们建议采用 SRMR, 及 NNFI、IFI、RNI、CFI、Gamma Hat、Mc、RMSEA 7 个指数中的一个组合来判断模型拟合, 这称为 2 指标策略。共有 SRMR + NNFI、SRMR + IFI、SRMR + RNI、SRMR + CFI、SRMR + Gamma Hat、SRMR + Mc、SRMR + RMSEA 7 个 2 指标策略<sup>[10, 11]</sup>。

### 1.3 关于 Hu 与 Bentler 结论的争议

Hu 与 Bentler 的研究设计严密, 其结论随后被相当多的人采纳<sup>[12]</sup>。但 Marsh 等通过模拟研究发现 Hu 与 Bentler 的判断标准有时会导致错误的结论, 因此不能简单地根据拟合指数判断模型是否拟合, 而且 Hu 与 Bentler 的界值标准定得过高, 不能过分推广<sup>[12]</sup>。温忠麟等重复了 Hu 和 Bentler 的部分模拟研究, 发现依据模型  $\chi^2$  所做出的判断要优于 2 指标策略<sup>[13]</sup>。Yuan 研究了不同研究条件(分布和样本量等)对拟合指数的影响, 结果发现所有拟合统计量都随样本量的变化而变化。因此 Yuan 认为拟合指数临界值标准是很难确定的, 但它们仍然是有意义的, 可用来判断模型的相对拟合<sup>[14]</sup>。Sivo 等发现在选择不同模型时拟合指数基本是稳定的, 而且与 Hu 与 Bentler 所报告的结果基本相同, 但 Sivo 等认为在不同样本量条件下都采纳 Hu 与 Bentler 的结论是不正确的, 他们认为还是应该为拟合指数确定一个标准<sup>[15]</sup>。

### 1.4 Hu 与 Bentler 研究设计的不足

在 Hu & Bentler 的模拟研究中设计了两个 CFA

<sup>\*</sup> 基金项目 辽宁师范大学省心理学重点实验室开放基金资助。

模型。一个被称为“简单”模型:15 个题目,每 5 个题测量一个因子,真模型中 3 个因子间存在相关,在错误模型中则有一个因子间相关被设定为 0。另一个被称为“复杂”模型:与“简单”模型相比增加了 3 个跨因子的负荷,错误模型中有一个跨因子载荷被设定为 0。Hu & Bentler 设计了 7 种不同的分布类型和 6 种样本容量。

Hu & Bentler 和温忠麟等都采用了这样的设计。但这种设计的问题是与研究实际有不小的差距,即研究结论不易概化。1)因子载荷过高,都在 0.70 以上 2)只考虑了参数少于真模型的参数不足模型,而没有考虑参数冗余模型和载荷(即参数)误置模型。而参数误置模型在实践中更为常见,因为在实际研究中对因子间是否有相关一般也不太关注,也不常遇到一个题目测量两个因素的情况。参数误置模型并不减少或增加自由度,在统计学上也更有意义。

研究拟借助模拟研究探讨 Hu 和 Bentler 推荐的拟合指数在更一般的条件下是否仍然能鉴别错误模型和正确模型,其 2 指标策略的效果是否好于单指标策略等。从而为结构方程模型拟合的判断提供更多的参考依据。

2 方法

采用如下设计:15 个题目(测量 3 个因素,每因素 5 个题),0.3、0.4、0.5、0.6、0.7、0.8 六种因子载荷,150、250、500、1000、2000、5000 六种样本容量。设计时只考虑观测变量为正态分布的情况。所有模型的因子间相关都是 0。Monte Carlo 数据模拟由

SAS9.1 软件完成,先根据理论模型生成符合多元正态分布的连续变量,再将数据取整转换成 7 点量表分数。

这样共有  $(\alpha \text{ 载荷量}) \times (\text{样本容量 } N) = 36$  种研究条件,对每种研究条件模拟出 200 个数据样本。对每个样本都设计一个真模型和一个参数误置(真模型中每个因子各由 5 个题目来测量,错误模型中则是有一个因子由 6 个题测量,另两个因子各由 4 个和 5 个题来测量)模型,用 LISREL8.53 进行模型估计,然后调整各拟合指标的临界值,探讨临界值为多大时对误设模型有最大的鉴别力。

3 结果

研究主要考察 Hu 和 Bentler 推荐的 8 个指标。首先考察它们取多大临界值时(即使用单指标时) $\alpha$  和  $\beta$  错误率最低,再考察其 2 指标策略。

3.1 以单个指标判断时的两类错误率

计算出指数 NNFI、IFI、RNI、CFI、Gamma Hat、Mc (RNI 与 CFI 性能相近,因此只考虑 CFI)在取临界值 0.85 ~ 0.99, RMSEA、SRMR 在取临界值 0.002 ~ 0.10 时的两类错误率之和。

结果发现 NNFI、IFI、CFI、Gamma Hat、Mc 的临界值定得越高,则对错误模型的拒绝率越高( $\beta$  错误越小),但对真模型的拒绝率也会增高( $\alpha$  错误变大)。当取某一临界值时  $\alpha$  与  $\beta$  错误率之和最小,则这一临界值就是最佳的界值。为节省篇幅,只给出了不同条件下各拟合指数的最佳界值及相应的两类错误率(见表 1)。

表 1 两类错误率之和最小时(%)各拟合指数的临界值

	NNFI	CFI	IFIP	Mc	G-hat	RMSEA	SRMR
N=150							
最佳界值	0.95	0.96	0.96	0.96	0.98	0.04	0.07
$\alpha + \beta$	24.2 + 17.5 = 41.7	24.2 + 17.2 = 41.6	23.6 + 17.9 = 41.5	26.2 + 26.1 = 52.3	8.3 + 39 = 47.3	5.5 + 43.1 = 48.6	8.5 + 22.5 = 31
N=250							
最佳界值	0.95	0.97	0.96	0.96	0.99	0.03	0.06
$\alpha + \beta$	13.6 + 15.8 = 29.1	18 + 13.3 = 31.3	14.6 + 16 = 30.6	8.4 + 29.4 = 37.8	10.2 + 28.3 = 38.5	6.1 + 32.3 = 38.4	0.3 + 32.7 = 33
N=500							
最佳界值	0.95	0.96	0.96	0.97	0.99	0.03	0.04
$\alpha + \beta$	8.6 + 10.5 = 19.1	8.9 + 10.2 = 19.1	8.4 + 10.8 = 19.2	2.4 + 25.9 = 28.3	0.6 + 29.6 = 30.2	0.3 + 31.7 = 32	4.5 + 12.7 = 17.2
N=1000							
最佳界值	0.94	0.95	0.95	0.99	0.99	0.02	0.03
$\alpha + \beta$	3.8 + 6.1 = 9.9	3.6 + 6.2 = 9.8	3.3 + 6.4 = 9.7	6.6 + 11.2 = 17.8	0 + 2.08 = 20.8	0.1 + 17.9 = 18	0.6 + 12.2 = 12.8

表 1  两类错误率之和最小时( % )各拟合指数的临界值

	NNFI	CFI	IFI	Mc	G-hat	RMSEA	SRMR
N = 2000							
最佳界值	0.95	0.96	0.95	0.99	0.99	0.01	0.02
$\alpha + \beta$	1.3 + 1.3 = 2.6	1.4 + 1.1 = 2.5	0.8 + 1.8 = 2.6	0.8 + 8.9 = 9.7	0 + 32.4 = 32.41	0.1 + 1.3 = 11.4	4.4 + 21.2 = 25.6
N = 5000							
最佳界值	0.93	0.95	0.95	0.99	0.99	0.01	0.02
$\alpha + \beta$	0 + 0.3 = 0.3	0 + 0.3 = 0.3	0 + 0.3 = 0.3	0 + 7.8 = 7.8	0 + 33.3 = 33.3	0.2 + 0.3 = 0.5	0 + 14.6 = 14.6

从表 1 中可以看出 7 个拟合指数在样本量不同时表现很不一致。与其他 4 个指数相比 ,NNFI、CFI、IFI 的表现是相对稳定的 ,即都是在 0.95 附近的两类错误率之和最小 ,且两类错误率之和比其他指数都小 ,因此这 3 个指数可视为优良的拟合指数。7 个指数中表现最差的是 Gamma Hat ,在所有样本容量条件下其两类错误率之和都大于 10% ,因此单独使用时要算性能很差的拟合指数。Mc 与 RMSEA 在 N 很大时才会有较低的两类错误率 ,且要求苛刻的界值标准 ,其性能也不理想。SRMR 的表现也不好。

从表中还可以看出 :1)在  $N \geq 1000$  时 ,如果 NNFI、CFI、IFI 中的一个取值在 0.95 左右时即可判定模型是拟合的 ,随着 N 的增大 ,0.95 的标准可适当放宽 ;2)在  $N < 1000$  时 ,即使 NNFI、CFI、IFI 中的一个取值在 0.95 左右也不能判断模型是拟合的 ,因为在这种条件下会有很高的  $\beta$  错误率。即使把标准提得很高 ,也不会把  $\beta$  错误率降低到 5% 以下。而根据其他 4 个指数做出判断时的效果更差。因此有必要考虑两个拟合指数相结合的策略 ,即 2 指标策略。

3.2 根据 Hu 和 Bentler 的 2 指标策略所做的判断

3.2.1 SRMR 与 NNFI、CFI、IFI、Mc、Gamma Ha 组合时

计算出 SRMR 取界值 0.03 ~ 0.10 ,同时 NNFI、CFI、IFI、Mc、Gamma Hat 中的一个取界值 0.90 ~ 0.97 时的两类错误率 ,结果为(表略) :1)在  $N = 150$  的条件下 :若 NNFI、CFI、IFI、Mc、Gamma Hat 中的一个取值大于 0.90 ,且  $SRMR < 0.06$  时 ,即有足够的信心判断模型是拟合的 ,因为这时接受错误模型的概率(  $\beta$  错误率 )仅为 1% ,但对真模型的拒绝率在 50 ~ 55% 之间。说明高于以上标准时可判断模型是拟合的 ,而低于以上标准时并不能认为模型错误。2)在  $N = 250$  的条件下 :NNFI、CFI、IFI、Mc、Gamma Hat 中的一个取值大于 0.90 ,且  $SRMR < 0.05$  时 ,才可判断模型是拟合的 ,但这时对真模型的拒绝率也很高( 33% )。3)在  $N = 500$  的条件下 :NNFI、CFI、IFI、Mc、Gamma Hat 大于 0.90 ,SRMR 的界值  $< 0.04$  时 ,则  $\alpha$

错误率都不高于 7% , $\beta$  错误率都不高于 13% ,这时也可初步判定模型是拟合的。4)在  $N \geq 1000$  的条件下 :若  $SRMR < 0.06$  或 0.05 时 ,则 NNFI、CFI、IFI 的值可放宽到 0.93 或 0.94 ;若  $SRMR < 0.03$  或 0.04 时 ,则 NNFI、CFI、IFI 的值可放宽到 0.91 或 0.92 ,在  $N > 1000$  时甚至可放宽到 0.90。

SRMR 与 Mc、Gamma Hat 组合时 ,即使将 Mc、Gamma Hat 的界值提高到 0.99 ,在 SRMR 的任何水平上都得不到满意的结果 ,且在 N 为 2000、5000 时也没有太多改善 ,因此在下文中将不再讨论这两个指数。

3.2.2 SRMR 与 RMSEA 组合时

让 SRMR 取界值 0.03 ~ 0.10 ,同时让 RMSEA 在取界值 0.01 ~ 0.10 ,计算两类错误率 ,得到的结果为 :

在  $N = 150$  时 ,SRMR 取界值 0.07 ,RMSEA 取界值 0.08 时两类错误率之和最小(  $\alpha = 8.6\%$  , $\beta = 22.5\%$  ) ;在  $N = 500$  时 ,SRMR 取界值 0.06 ,RMSEA 取界值 0.08 时两类错误率之和最小(  $\alpha = 0.25\%$  , $\beta = 32.7\%$  ) ;在  $N = 500$  时 ,SRMR 取界值 0.04 ,RMSEA 取界值 0.08 时两类错误率之和最小(  $\alpha = 4.5\%$  , $\beta = 12.7\%$  )。因此在  $N = 150 \sim 500$  之间时 ,如果 SRMR 与 RMSEA 都不大于以上所列的相应界值 ,即可初步判断模型是拟合的。

在  $N = 1000$  时 ,SRMR 取界值 0.05 ,RMSEA 取界值 0.02 时两类错误率之和为 : $\alpha = 1\%$  , $\beta = 14.9\%$  ;SRMR 取界值 0.03 ,RMSEA 取界值 0.09 时两类错误率之和为 : $\alpha = 0$  , $\beta = 16.9\%$ 。在  $N = 2000$  时 ,SRMR 取界值 0.05 ,RMSEA 取界值 0.02 时两类错误率之和为 : $\alpha = 0.2\%$  , $\beta = 16.3\%$  ,SRMR 取界值 0.02 ,RMSEA 取界值 0.09 时两类错误率之和为 : $\alpha = 4.4\%$  , $\beta = 0.4\%$ 。在  $N = 5000$  时 ,SRMR 取界值 0.05 ,RMSEA 取界值 0.02 时两类错误率之和为 : $\alpha = 0.2\%$  , $\beta = 16.3\%$  ,SRMR 取界值 0.02 ,RMSEA 取界值 0.09 时两类错误率之和为 : $\alpha = 4.4\%$  , $\beta = 0.4\%$ 。因此在

$N = 1000 \sim 5000$  的条件下如果 SRMR 与 RMSEA 都不大于以上所列的相应界值,即可判断模型是拟合的。

RMSEA 与 SRMR 组合时有时也能得到较满意的两类错误率,但仍大于 NNFI、CFI、IFI 与 SRMR 的组合结果。

### 3.3 单指标与 2 指标策略所得到的两类错误之和的最小值

对比使用单个拟合指数和 2 指标策略做出判断时的两类错误之和的最小值(表略),可以看出绝大多数情况下都是 2 指标策略的两类错误之和更小。这说明使用 2 指标策略确有其合理性。使用 2 指标策略可以说主要有两点好处,一是降低两类错误率,二是使单个指数的界值不至定得太高。

## 4 讨论与结论

研究采用了与 Hu & Bentler 不同的设计,得出的结果有的与他们有的一致而有的不一致。1)支持 2 指标策略的有效性。使用 2 指标策略确实能降低两类错误的发生概率,原因是不同指数对不同类型的模型错误敏感度不同。2)使用单个拟合指数时 NNFI、CFI、IFI 的界值在 0.95 左右比较好,但 RMSEA 和 SRMR 的最佳界值则比他们建议的还要低。3)Mc 和 Gamma-Hat 表现不好,不能算作优良的指数。因此不建议使用。NNFI、CFI、IFI 的性能相当好,应算优良的指数,SRMR 单独使用时性能并不太好,但与其他指标一起使用时确能增加对错误模型的敏感性。4)确定界值标准时要考虑样本容量。若单独使用 NNFI、CFI、IFI,则在不同  $N$  条件下界值定为 0.95 左右都是最好的,但 RMSEA 和 SRMR 则不然,在使用 2 指标策略时界值随  $N$  的变化则更大。因此建议研究者根据样本量确定拟合指数的界值。

综合分析模拟研究结果,可得到以下结论:1)在  $N \geq 1000$  时,只要 NNFI、CFI、IFI 中的一个取值在 0.95 左右即可判断模型是拟合的,这时不需要使用 2 指标策略,也不需要报告其他的指数。2) $N \geq 1000$ ,但 NNFI、CFI、IFI 的取值都低于 0.95,模型是否拟合就取决于 SRMR。若 SRMR 的值低于 0.06 或 0.05 则 NNFI、CFI、IFI 的值可放宽到 0.93 或 0.94;若 SRMR 的值低于 0.03 或 0.04 时,则 NNFI、CFI、IFI 的值可放宽到 0.91 或 0.92。在  $N$  特别大时 NNFI、CFI、IFI 的界值还可适当降低,但只有 NNFI 可放宽到 0.90。3)在  $N \leq 500$  时即使 NNFI、CFI、IFI 三个指标中有一个大于 0.95 仍没有足够大的把握判断模型是拟合的,因为这时接受错误模型的概率都大于

10%,在  $N$  接近 500 时也只能初步判断模型是拟合的。4)在  $N \leq 500$  可采用 NNFI、CFI、IFI 与 SRMR 结合来判断模型是否拟合,以降低两类错误率。若 NNFI、CFI、IFI 中的一个取值在 0.90 左右,要判断模型是拟合的,则在  $N = 150$  时,SRMR 需  $\leq 0.06$ , $N = 250$  时,SRMR 需  $\leq 0.05$ , $N = 500$  时,SRMR 需  $\leq 0.04$ 。

研究所使用的统计软件是 LISREL8.53,而 Hu & Bentler 使用的则是 LISREL8.2。8.52 以后版本所计算出的 CFI、NNFI 值增大,其他指标基本不变<sup>[16]</sup>。这一增加量大约是 0.03 左右。因此如果研究者使用的是 8.52 以后的版本,在判断模型拟合时参照上述标准即可,但若是 8.52 以前的版本,则 CFI、NNFI 等的值应减少 0.03。

以上结论是根据模拟数据得出的。在真实数据中情况要复杂得多,比如题目分数不符合正态分布,一个题目测量两个或多个因素,题目的误差项间存在相关等。研究者应谨慎参考研究的结论,并适当将拟合指数的界值放宽。

结构方程模型拟合的判断确实是个非常复杂的问题。样本量不同,待估参数数量不同时拟合指数的界值往往也应有所不同。除使用性能良好的拟合指数外,修正指数也是一个重要的指标,修正指数很大时表明模型中有错误设定的路径。因此除考虑拟合指数外还要根据修正指数修改研究模型,在修正指数都比较小时才说明模型是完善的。

针对研究发现,建议在验证性因素分析时尽量使用大的样本量(样本量相当于题目量的 10 倍以上),另外,凭一次验证性因素分析的结果可能并不能确保研究模型的正确性,研究者最好能做交叉样本的检验。

## 参考文献

- 1 Jöreskog K G, Sörbom D. LISREL 8: User's reference guide. Chicago: IL Scientific Software international, 1996.
- 2 Hu L T, Bentler P. Evaluating model fit. In: R. H. Hoyle, Ed. Structural Equation Modeling: Concepts, Issues, and Applications. London: Sage, 1995. 76 - 99.
- 3 Bentler P M. EQS structural equations program manual. Encino, CA: Multivariate Software, 1995.
- 4 Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Structural Equation Modeling, 1999 (6) 56 - 83.
- 5 Muthén L K, Muthén B O. Mplus User's Guide. Los Angeles, CA, 2001.
- 6 Bollen K A. Overall fit in covariance structure models: Two types

- of sample size effects. *Psychological Bulletin*, 1990, 107 :256 – 259.
- 7 Marsh H W ,Balla J R ,McDonald R P. Goodness – of – fit indexes in confirmatory factor analysis :The effect of sample size. *Psychological Bulletin* ,1988 ,103 ( 3 ) 391 – 410.
  - 8 McDonald R P ,Marsh H W. Choosing a multivariate model :Non – centrality and goodness – of – fit. *Psychological Bulletin* , 1990 ,107 247 – 255.
  - 9 Marsh H W ,Hau K – T. Assessing goodness of fit :Is parsimony always desirable ? *The Journal of Experimental Education* ,64 , 1996 364 – 390.
  - 10 Hu L ,Bentler P M. Fit indices in covariance structure analysis : Sensitivity to underparameterized model misspecification. *Psychological Methods* ,1998 ( 3 ) 424 – 453.
  - 11 Hu L ,Bentler P M. Cutoff criteria for fit indexes in covariance structure analysis :Conventional criteria versus new alternatives. *Structural Equation Modeling* ,1999 6 ( 1 ) :1 – 55.
  - 12 Marsh H W ,Hau K – T ,Wen Z. In :search of golden rules : Comment on hypothesis testing approaches to setting cut off values for fit indexes and dangers in overgeneralising Hu & Bentler's ( 1999 ) findings. *Structural Equation Modeling* ,2004 , 11 ( 3 ) 320 – 341.
  - 13 温忠麟 ,侯杰泰 ,马什赫伯特 . 结构方程模型检验 拟合指数与卡方准则 . *心理学报* 2005 36 ( 2 ) :186 – 194.
  - 14 Yuan K – H. Fit Indices Versus Test Statistics. *Multivariate Behavioral Research* 2005 40 ( 1 ) :115 – 148.
  - 15 Sivo S A ,Fan X ,Willse J T. The search for “ Optimal ” Cutoff Properties :Fit Index Criteria in Structural Equation Modeling. *The Journal of Experimental Education* ,2006 ,74 ( 3 ) :267 – 288.
  - 16 Schumke S C ,Hardt J. A cautionary note on Incremental Fit Indices reported by LISREL. *Methodology* ,2005 ,1 ( 2 ) :81 – 85.

## Assessing the Model Fit in Confirmatory Factor Analysis

Guo Qingke Wang Weili Chen Xuexia

( Liaoning Normal University ,Department of Psychology ,Dalian 116029 )

**Abstract** :In this simulation study ,we designed 6 loading size conditions ( 0.03 ~ 0.08 ) and 6 sample size conditions ( 150 ,250 ,500 ,1000 , 2000 ,5000 ) to test the capacity of the 8 fit indices and 2 – index strategies recommended by Hu & Bentler to avoid  $\alpha$  and  $\beta$  type error. The true model is also a CFA model with 3 factors and 15 items ,but each item only loads on one factor ,and in the wrong model there is a mis – specified factor loading. The results show that among the 8 fit indices NNFI ,CFI and IFI always produce the least sum of  $\alpha$  and  $\beta$  type error ,their best cut – off values consistently fall around 0.95 across all sample size conditions. 2 – index strategies can produce less sum of  $\alpha$  and  $\beta$  type error compared to single index strategies ,they are especially useful when sample size is small.

**Key words** Structural Equation Modeling ,Model fit ,Fit Indices , $\alpha$  and  $\beta$  type error