

• 研究方法(Research Method) •

惩罚对齐法：测量不变性检验的新方法

温聪聪

(厦门大学国际学院, 厦门 361005)

摘要 Asparouhov 和 Muthén 在 2023 年提出了一种全新的惩罚结构方程模型框架。惩罚对齐法是该模型框架在测量不变性检验领域的应用范例。惩罚对齐法继承了多组探索性因子分析在探索性因子分析框架内进行测量不变性检验和可以估计交叉载荷等优点, 继承了对齐法使用成分损失函数、允许模型中存在一定量不等参数等优点, 同时继承了贝叶斯结构方程模型对模型参数设置先验分布、通过检验模型参数的近似测量不变性达到潜因子均值比较的目的等优点。此外, 惩罚对齐法还克服了传统测量不变性检验方法的一些不足。本文以大学生职业价值观研究为例, 比较了多组验证性因子分析、基于验证性因子分析的惩罚对齐法分析、基于探索性结构方程模型的惩罚对齐法分析拟合样本数据的效果, 演示了如何使用惩罚对齐法进行测量不变性检验和多组比较。

关键词 测量不变性检验, 惩罚对齐法, 贝叶斯结构方程模型, 探索性结构方程模型, 对齐先验分布

分类号 B841

1 研究背景

在社会科学实证研究中进行各类测量评估的时候, 往往涉及多个群体的多组比较。多组验证性因子分析就是一种进行多组比较的传统统计方法。通过该方法对各组的潜因子均值进行直接比较之前, 需要检验模型的测量不变性(白新文, 陈毅文, 2004; 刘军, 吴维库, 2005; Asparouhov & Muthén, 2014; Millsap, 2012)。一般情况下, 当模型满足因子载荷和截距严格相等的强测量不变性时, 才能对各组的潜因子均值进行直接比较(王孟成, 2014; 王阳 等, 2022; 温聪聪 等, 2019)。但因因子载荷和截距严格相等在模型具有较多组数或模型较复杂等条件下很难满足, 导致多组比较无法进行(宋琼雅 等, 2021; 王阳 等, 2020; 温聪聪, 史秋衡, 2022; Asparouhov & Muthén, 2023a)。

为了解决这一问题, 研究者们最初依靠模型修正指数来修正模型, 使模型满足部分测量不变性。模型修正指数代表的是将模型中的一个限制

参数变为自由估计, 所减少的 1 个自由度对应的卡方统计量减小值。当限制参数的修正指数较大时, 说明需要将其改设为自由估计, 逐渐修正模型。但由于修正指数会受到样本量的影响, 因此很难有统一的修正指数达到多大时需要将限制参数改设为自由估计的标准。因此, 依据修正指数来修正模型往往具有一定的主观性。在实际操作中, 模型如果越复杂, 修正过程也越复杂, 主观决定的模型修正过程以及修正指数间存在的多重共线性问题可能导致修正出的模型偏离真实正确的模型, 潜因子均值的估计结果很可能出现偏差(Asparouhov & Muthén, 2014; Marsh et al., 2018)。

近 10 年来, 研究者们提出了不少前沿的测量不变性检验方法, 例如贝叶斯结构方程模型(Muthén & Asparouhov, 2013)、等效性检验(Deng & Yuan, 2016)、对齐探索性结构方程模型(Asparouhov & Muthén, 2023a)、二水平随机效应法(Muthén & Asparouhov, 2018)。这些方法不同程度地解决了传统测量不变性检验方法的不足, 为更加科学地进行多组比较提供了更多选择。不过, 多数方法的适用情境有限, 并且都有各自的不足。2023 年, Asparouhov 和 Muthén 提出了一种

收稿日期: 2023-08-01

通信作者: 温聪聪, E-mail: wencong001@xmu.edu.cn

惩罚结构方程模型(Penalized structural equation model, PSEM)框架。该框架在对齐法成分损失函数的原理基础上进一步推广,使惩罚函数可以依据不同的研究情境自设,拥有较广泛的适用性(Asparouhov & Muthén, 2023b)。同时,依据该模型框架开发的惩罚对齐法整合了多组探索性因子分析、对齐法、贝叶斯结构方程模型等方法的优点,并且克服了这些方法的不足。在 Mplus 最新的 8.10 版本软件中,已经可以使用惩罚对齐法进行测量不变性检验。惩罚对齐法不仅可以应用于多组比较,还可以应用于跨时序纵贯数据分析。

本研究首先对测量不变性检验方法近十几年来的发展历史和较为前沿的几种测量不变性检验方法进行回顾,指出每个方法的优点和不足。其次,惩罚对齐法是惩罚结构方程模型框架应用于测量不变性检验的范例,本研究详述其运算过程,指出惩罚对齐法继承了传统测量不变性检验方法的哪些优势,克服了哪些不足。最后,运用一则大学生职业价值观研究实例演示如何使用惩罚对齐法进行测量不变性检验和多组比较。

2 测量不变性检验方法进展回顾

测量不变性检验在社会科学实证研究中得到了广泛使用,但由于因子载荷和截距跨组严格相等的假设较为苛刻,学界在不断研发比该假设限制更弱的统计方法。Asparouhov 和 Muthén(2009)首先研发出了探索性结构方程模型(Exploratory Structural Equation Model, ESEM),应用于测量不变性检验时被称为多组探索性因子分析。多组探索性因子分析具有不必预设因子载荷矩阵结构,可准确估计模型中的中低程度交叉载荷的优点。Muthén 和 Asparouhov (2013)紧接着又将贝叶斯结构方程模型(Bayesian Structural Equation Model, BSEM)应用于测量不变性检验,首次将检验模型中因子载荷和截距跨组严格相等的假设转变为检验因子载荷和截距跨组近似相等。Jak 等人(2013)将多水平分析模型框架应用于测量不变性检验,该方法被称为二水平随机效应法(two-level random effects model, TREM)。二水平随机效应法将个案间的差异分离为组内差异和组间差异(Muthén & Asparouhov, 2018)。

贝叶斯结构方程模型、二水平随机效应法等检验模型近似测量不变性的方法在不同程度上减

弱了潜因子均值比较需要模型中的因子载荷和截距严格相等的前提条件限制,多组探索性结构方程模型也解决了交叉载荷的估计问题。但有一个问题始终无法避开:当模型中存在一定量的不等参数,且这些不等参数的不等程度较高,明显不满足近似测量不变性时,ESEM、BSEM、TREM 的分析效果普遍不尽如人意(Asparouhov & Muthén, 2014; Asparouhov & Muthén, 2023a)。

Asparouhov 和 Muthén (2014)提出了一种全新的对齐法(Alignment),利用类似探索性因子分析中旋转函数的成分损失函数 $f(x) \approx \sqrt{|x|}$ 对模型参数的解进行优化,识别原本多组验证性因子分析形态不变模型无法识别的潜因子方差和潜因子均值,使模型具有少量较高程度不等参数和大多数满足近似测量不变性的参数。与 ESEM、BSEM、TREM 等方法的一个重要不同是对齐法是基于模型拟合度较好的多组验证性因子分析的形态不变模型进行优化计算,而其他方法是检验因子载荷和截距严格相等或近似严格相等模型。该方法是测量不变性检验领域的一个重要进展,使多组比较可以绕开检验因子载荷和截距严格相等的假设进行。Asparouhov 和 Muthén (2023a)进一步提出了对齐探索性结构方程模型(Aligned-Exploratory Structural Equation Model, AESEM),该模型将对齐法和探索性结构方程模型的优势整合,可以有效估计模型中的不等参数、交叉载荷、潜因子间协方差等,克服了传统对齐法的不足。随后,Asparouhov 和 Muthén (2023b)进一步提出了惩罚结构方程模型(PSEM)框架,本文要详细介绍的惩罚对齐法就是将惩罚结构方程模型框架应用于测量不变性检验的范例。

国内心理统计与测量领域的研究者们与时俱进,发表了一些关于测量不变性检验方法的研究,独创出新的测量不变性检验方法或是将国外前沿测量不变性检验方法引入我国。张沥今等人(2019)、宋琼雅等人(2021)研究了使用贝叶斯结构方程模型进行测量不变性检验和多组比较的步骤和注意事项。温聪聪通过蒙特卡洛模拟研究探究了对齐法在单因子和三因子模型中允许的不等参数率(温聪聪,史秋衡,2022),进一步探究了对齐法的适用性。Deng 和 Yuan (2016)首次提出了投影法,绕过因子载荷和截距严格相等的假设,基于因子载荷严格相等模型进行等效性检验以达到潜

因子均值比较的目的。王阳等人(2022)总结了结构方程模型的参数估计方法、拟合优度评价指标、特殊数据的处理,测量不变性检验的几种方法及其适用情境等内容,回顾了新世纪20年以来我国心理统计与测量领域结构方程模型建模和测量不变性检验等方面的研究进展。以上这些研究提出、引入或是总结了测量不变性检验方法的新进展,使我国实证研究者在进行测量不变性检验时有了更多选择,也更加了解如何使用贝叶斯结构方程模型、对齐法和投影法等方法进行测量不变性检验和多组比较。

以下,本文将回顾目前较为前沿的几种测量不变性检验方法,指出每个方法所能解决的问题,以及它们的优点和不足。

2.1 贝叶斯结构方程模型(BSEM)

贝叶斯估计法在统计学领域属于主流方法,其成为主流的一个重要原因是马尔科夫链蒙特卡罗算法取得了巨大成功。贝叶斯结构方程模型是检验模型测量不变性的方法之一。因子载荷和截距严格相等的假设过于苛刻,现实中很难满足(宋琼雅等, 2021; 王阳等, 2020; 温聪聪, 史秋衡, 2022; Asparouhov & Muthén, 2023a; Muthén & Asparouhov, 2013),而基于贝叶斯结构方程模型进行测量不变性检验的优势则是将属于不同组的两两对应的因子载荷参数和截距参数之间的差异看作待估计参数,对这些参数设置先验分布。这些先验信息设定了在分析实证数据获得的信息前提下的条件概率分布中不确定性的多少,反映了研究者对模型的已知信息或是先前理论提供的信息。先验分布可以依据数据分析的实际情况设置信息先验分布(informative prior)和无信息先验分布(noninformative prior)。如果对待估计参数所知信息较少,则可以为该参数的先验分布设置一个较大方差,体现出参数值较大的不确定性(Muthén & Asparouhov, 2012)。相反,如果对参数值有较为明确的把握,可以确定是在某个小的范围内变动,则可以为该参数的先验分布设置一个较小的方差(王孟成等, 2017; 张沥今等, 2019; Muthén & Asparouhov, 2012)。

因此,进行单组贝叶斯结构方程模型检验不再需要将接近为0的交叉载荷设置为零载荷,而是需要设为近似零载荷,即通过设置正态先验分布的0均值和先验方差预测其估计值的范围。在

使用极大似然估计法的验证性因子分析中,将交叉载荷引入模型会导致模型无法识别。而贝叶斯结构方程模型引入小先验分布方差可以为模型估计提供更多信息,避免无法识别问题。一般情况下,先为这些交叉载荷设置0均值和0.01的先验方差,则这些交叉载荷估计值的95%可信区间(Credible Intervals)应落在 $[-0.196, 0.196]$,意味着低程度的交叉载荷,接近0但不为0。0.01这样较小的先验方差能让参数估计值的可信区间不偏离0过多,避免得出错误的PPp值(Muthén & Asparouhov, 2012)。设置0.01的先验方差运行出结果后,如果交叉载荷估计值的可信区间(Credible Intervals)明显没有落在 $[-0.196, 0.196]$ 的区间里,那么就需要依据估计值大小增大这些交叉载荷的先验方差,再次运行新模型拟合数据。使用贝叶斯结构方程模型进行多组测量不变性检验也不再要求模型中的所有因子载荷和截距参数跨组严格相等,而是将跨组严格相等的参数设置为近似严格相等,将跨组参数间的差异设置信息先验分布或无信息先验分布,模型最终满足近似测量不变性即可(Muthén & Asparouhov, 2013)。这一特点使贝叶斯结构方程模型具有处理一定量低程度交叉载荷和低程度不等参数的能力。此外,贝叶斯结构方程模型还可以引入残差相关,为进一步提升模型的正确性提供了可能(Muthén & Asparouhov, 2012; Muthén & Asparouhov, 2013)。

2.2 等效性检验(Equivalence testing)

多组验证性因子分析和多组探索性结构方程模型等测量不变性检验方法在判定嵌套模型优劣时是通过似然比零假设检验来进行。在参数满足多元正态分布和嵌套模型成立的假设前提下,嵌套模型和基线模型的似然比卡方变化值渐近服从中心卡方分布,将卡方变化值、对应的自由度变化值相结合计算出对应 p 值,与显著性水平 α 比较后得出卡方变化统计量的显著性,判断测量不变性程度更强的嵌套模型是否被拒绝。似然比零假设检验有两个不足之处,一是零假设没有被拒绝有可能只是说明拒绝零假设的效力较低,也就是统计学上的二类错误,并不一定说明零假设成立(Shi et al., 2019)。二是似然比零假设检验探究的是因子载荷严格相等或因子载荷和截距严格相等的嵌套模型是否被拒绝。在极大似然估计中,

当数据样本量较大时,即使一些因子载荷或截距的不等程度很微小,但会统计显著。在许多高组数目、高样本量的跨国、跨文化多组比较当中,因子载荷严格相等的假设和因子载荷和截距严格相等的假设都很容易被拒绝(Byrne & van de Vijver, 2010; Caycho-Rodríguez et al., 2021; Jang et al., 2017; Lomazzi, 2018; Munck et al., 2018)。综上所述,传统测量不变性检验方法采用似然比零假设检验存在一定的局限。

等效性检验在这一背景下被提出。其计算逻辑是设定测量不变性程度较强的嵌套模型,计算出嵌套模型成立对应的似然比统计量,并且将似然比进一步转化为 RMSEA 或 CFI 模型拟合指数。根据先前研究, RMSEA 和 CFI 指标表现的好坏已经有了临界值标准,将转化出的模型拟合指数值与临界值标准相比较,就能判断该嵌套模型是否被拒绝(Yuan et al., 2016)。

Deng 和 Yuan (2016)基于等效性检验进一步提出了投影法(projection method)用于测量不变性检验。该方法将每组中每个观测指标的均值分为两个相互正交的成分,即共同分数和特殊因子。检验共同分数均值的跨组不变性和传统方法检验潜因子均值的跨组不变性是等效的。使用该方法比较各组的潜因子均值不必满足因子载荷和截距严格相等的假设,只需要满足因子载荷严格相等的假设,直接检验共同分数均值的跨组不变性即可。这使测量不变性检验变得更加容易(王阳等, 2020; Deng & Yuan, 2016)。

虽然投影法和等效性检验为测量不变性检验提供了新的可能和做法,但它们也存在一些不足。第一,当结果拒绝零假设时,只能说明已有检验结果和指标不支持嵌套模型的测量不变性,但并不能肯定模型参数的不等性被发现了(Shi et al., 2019)。第二,两种方法判断模型测量不变性的标准主要是基于 RMSEA 和 CFI 指标,但这些指标并不是标准化的指标,不具备比较性,提供不了更多有关模型参数不等程度的信息。例如,如果两个模型的 RMSEA 值都为 0.13,这一值大于 0.08 的可接受标准。虽然两个模型的 RMSEA 值相对可接受标准的差值都为 0.05,但两个模型的复杂程度、参数的不等性、样本量等基本条件可能有很大不同。第三,使用投影法进行潜因子均值比较需要模型满足弱测量不变性,但该假设在

一些高组数目、高样本量的多组比较模型中仍然很难满足(温聪聪, 史秋衡, 2022; Asparouhov & Muthén, 2023a)。正是因为这些不足,投影法和等效性检验并没有在多组比较研究中被广泛使用。

2.3 对齐法和对齐探索性结构方程模型(Alignment 和 AESEM)

传统方法进行潜因子均值多组比较都需要检验因子载荷和截距严格相等的假设。即使 Muthén 和 Asparouhov 将贝叶斯结构方程模型应用于多组比较,也需要通过设定正态先验分布检验模型的“强近似测量不变性”,然后由后往前修正模型,实现多组比较。2014 年,Asparouhov 和 Muthén 提出了对齐法,基于形态不变模型使用该方法进行测量不变性检验和多组比较。形态不变模型原本是自由估计因子载荷和截距,潜因子均值被设为 0,潜因子方差被设为 1。但在对齐法中,潜因子均值和潜因子方差也被自由估计。为了估计潜因子均值和潜因子方差需要依据以下总损失函数 F 进行优化计算,使原本无法识别的模型变得可识别(Asparouhov & Muthén, 2014)。

$$F = \sum_p \sum_{g^1 < g^2} W_{g^1, g^2} f(\lambda_{pg^1} - \lambda_{pg^2}) + \sum_p \sum_{g^1 < g^2} W_{g^1, g^2} f(\tau_{pg^1} - \tau_{pg^2}) \quad (1)$$

其中 p 为观测指标序数, g 为组序数, λ_{pg} 为因子载荷, τ_{pg} 为截距。 $W_{g^1, g^2} = \sqrt{N_{g^1} N_{g^2}}$, N_{g^1} 、 N_{g^2} 代表组样本量,样本量大的组对总损失函数的贡献要高于样本量小的组。 $f(\lambda_{pg^1} - \lambda_{pg^2})$ 和 $f(\tau_{pg^1} - \tau_{pg^2})$ 分别代表 g^1 组和 g^2 组因子载荷和截距的差异,被称为成分损失函数,如下式所示:

$$f(x) \approx \sqrt{|x|} \quad (2)$$

为了理清对齐法使用成分损失函数 $f(x) \approx \sqrt{|x|}$ 求最小化解的逻辑,使用如图 1 所示的 $f(x) \approx \sqrt{|x|} - x$ 的函数图来说明。当 $0 < x < 0.25$ 时, $\sqrt{|x|} > x$, 且 $\sqrt{|x|} - x$ 的值单调递增,成分损失函数 $f(x) \approx \sqrt{|x|}$ 被高估的程度越来越大;当 $0.25 < x < 1$ 时, $\sqrt{|x|} > x$, 且 $\sqrt{|x|} - x$ 的值单调递减,成分损失函数 $f(x) \approx \sqrt{|x|}$ 被高估的程度越来越小;当 $x > 1$ 时, $\sqrt{|x|} < x$, $f(x) = \sqrt{|x|} - x$ 为负数,成分损失函数 $f(x) \approx \sqrt{|x|}$ 被低估。所以,根据 $f(x) \approx \sqrt{|x|}$ 的性质,当 x 取值接近 0 或接近 1 时,

$\sqrt{|x|}-x$ 的高估程度很小; 当 x 取值大于 1 时, $f(x) \approx \sqrt{|x|}$ 被低估。所以, 总损失函数 F 取得最小值的解应是当模型拥有少量较大不等参数和大多数满足近似测量不变性参数的时候。

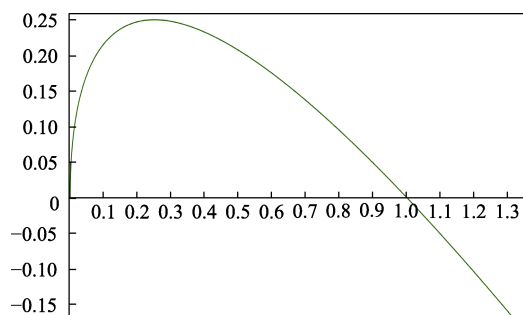


图 1 $f(x) = \sqrt{|x|} - x$ 在第一、四象限部分区间的函数图

对齐法的运算原理是对于任意一整套的潜因子均值和潜因子方差, 必然有相对应的截距和因子载荷, 使对齐模型和形态不变模型具有相等的似然值。利用成分损失函数 $f(x) \approx \sqrt{|x|}$ 的特点使模型在拥有少数较大不等参数和大部分满足近似测量不变性参数的时候使模型不等性最小化, 此时得出的模型就是最优对齐模型。根据蒙特卡洛模拟研究显示, 对齐法在多因子模型中存在 20% 至 30% 的不等参数时, 仍能提供精确的模型参数估计结果。当模型为单因子模型时, 对齐法允许模型中存在的不等参数率可以达到 100% (温聪聪, 史秋衡, 2022)。

当总损失函数 F 被最小化后, 对齐法通常可以识别所有组的潜因子均值和除第一组外所有组的潜因子方差。对齐法的参数优化方式可以分为固定对齐优化(fixed alignment optimization)和自由对齐优化(free alignment optimization)两种方式。固定对齐优化方式设定第一个组的潜因子均值 a_1 为定值 0, 该组也被称为参照组(参照组默认为 g_1 , Mplus 中可以自行设置改变参照组), 而自由对齐优化方式则设定 a_1 为自由估计参数。第一组的潜因子方差 ψ_1 则需要利用以下参数限定识别并估计:

$$\psi_1 \times \psi_2 \times \cdots \times \psi_g = 1 \quad (3)$$

为了使参数标准化, 两种算法在报告输出结果时将 ψ_1 转换为定值 1 的尺度来报告最终参数估计结果。

对齐法的提出使测量不变性检验不再依赖检验因子载荷和截距严格相等, 而是检验形态不变模型中因子载荷和截距参数是否近似严格相等。对齐法具有处理更多不等参数的能力, 同时可以为模型中的所有参数提供测量不变或不等的诊断。对齐法在大型跨国、跨文化研究中已经越来越被广泛使用(Byrne & van de Vijver, 2010; Caycho-Rodríguez et al., 2021; Jang et al., 2017; Lomazzi, 2018; Munck et al., 2018)。

虽然对齐法有这些独特的优势, 但该方法仍是基于传统的验证性因子分析模型进行分析, 无法在模型中引入交叉载荷, 这将导致主因子载荷、潜因子相关性的高估, 以及潜因子均值估计的偏差(Wen & Hu, 2022)。2023 年, Asparouhov 和 Muthén 将对齐法和探索性结构方程模型进行了整合, 进一步提出了对齐探索性结构方程模型(AESEM)。该方法的运算过程主要分为三个步骤: 首先使用未旋转模型拟合数据样本, 其次基于探索性因子分析的旋转标准对未旋转模型进行拟合, 最后基于对齐法成分损失函数对形态不变的探索性因子分析模型进行拟合(Asparouhov & Muthén, 2023a)。对齐探索性结构方程模型兼具探索性结构方程模型可以精确估计交叉载荷和对齐法允许模型中存在一定量不等参数的优势, 是基于近似测量不变性检验进行多组比较的一个十分有应用前景的新方法。

2.4 二水平随机效应法(two-level random effects model)

二水平随机效应法将测量不变性检验应用于多水平模型中, 使测量不变性检验拥有多水平模型具备的优势。根据 Muthén 和 Asparouhov (2018) 的模型表述, 对于一个特定的连续型潜因子观测指标 y_{ij} , 二水平随机效应模型如下式所示:

$$y_{ij} = v_j + \lambda_j f w_{ij} + \varepsilon_{ij} \quad (4)$$

$$v_j = v + \lambda f_{Bj} + \varepsilon v_j, \quad (5)$$

$$\lambda_j = \lambda + \lambda f_{\psi j} + \varepsilon \lambda_j. \quad (6)$$

其中 $f w_{ij} \sim N(0, 1)$, $\varepsilon_{ij} \sim N(0, \theta)$, $\varepsilon v_j \sim N(0, \sigma_v^2)$, $\varepsilon \lambda_j \sim N(0, \sigma_\lambda^2)$, $f_{Bj} \sim N(0, \psi)$, $f_{\psi j} \sim N(0, \sigma^2)$ 。 σ_v^2 体现的是截距的变异, σ_λ^2 体现的是因子载荷的变异, ψ 体现的是潜因子均值的变异, σ^2 体现的是潜因子方差的变异。组间水平的 f_{Bj} 潜因子得分均值就是各个组的潜因子均值。

使用二水平随机效应法进行分析通常需要遵

循 Jak 等人(2013)提出的三个随机截距模型分析步骤: (1)设定组内和组间因子载荷不等, 组间水平的残差方差自由估计; 这个模型把 σ^2 和 σ_{λ}^2 设为 0, 且令 $\lambda_j = \lambda w$, σ_v^2 自由估计。(2)组内和组间因子载荷相等, 组间水平的残差方差自由估计, 这一步设置了弱近似测量不变性; 这个模型把 σ^2 和 σ_{λ}^2 设为 0, 且令 $\lambda_j = \lambda$, σ_v^2 自由估计。(3)组内和组间因子载荷相等, 组间水平的残差方差为 0, 这一步设置了强近似测量不变性(Kim et al., 2017); 这个模型把 σ^2 和 σ_{λ}^2 设为 0, 且令 $\lambda_j = \lambda$, σ_v^2 也设为 0。

二水平随机效应法将模型参数看作是随机参数, 检验的是模型参数的近似测量不变性。由于多水平模型较为复杂, 使用极大似然估计法估计模型参数会使定积分的运算维度变多, 很难进行运算, 所以二水平随机效应法使用的是贝叶斯估计法。二水平随机效应法具备可以处理组数目超过 100 的模型, 组所需最小样本量要求较低, 可以设定组水平变量影响模型参数, 通过检验模型参数的近似测量不变性达到多组比较目的等优点(Muthén & Asparouhov, 2018)。

同时, 二水平随机效应法也存在一些不足, 例如该方法要求模型中的因子包含较多数量的观测指标, 测量参数必须满足正态分布假设, 在复杂抽样调查中较难适用, 运算时间较长等(Muthén & Asparouhov, 2018)。

已有的测量不变性检验方法还有多水平因子混合模型、多组探索性结构方程模型等其他方法。由于这些方法的适用性有限, 所以未纳入本研究的回顾中。

3 惩罚对齐法——惩罚结构方程模型(PSEM)框架在测量不变性检验领域的应用

惩罚结构方程模型于 2023 年由 Asparouhov 和 Muthén 提出, 该研究目前已可在 Mplus 官方网站阅读(Asparouhov & Muthén, 2023b)。惩罚结构方程模型可以应用于许多已有的统计分析方法, 包括跨时序结构方程模型参数的对齐(比如潜在增长曲线模型), 结构方程模型参数的对齐(比如对齐结构方程模型), 多组探索性因子分析(比如多组探索性结构方程模型和对齐探索性结构方程模型), 使用非正交和非斜交旋转标准的探索性因

子分析(比如二阶探索性因子分析和拥有超过一个主因子的双因子探索性因子分析), 所有组的因子均为正交关系的强测量不变多组探索性因子分析, 模型参数满足近似测量不变性或近似为零的贝叶斯结构方程模型等。由于只是使用了相应方法的旋转标准、成分损失函数等作为惩罚函数, 以上列出的情况和方法并未穷尽惩罚结构方程模型的适用情境(Asparouhov & Muthén, 2023b)。在实践中, 可以自设惩罚函数, 应用于新的模型和情境。

3.1 惩罚对齐法进行测量不变性检验的原理

惩罚结构方程模型框架应用于测量不变性检验和多组比较的一个典型案例是惩罚对齐法。接下来详述惩罚对齐法进行测量不变性检验的原理。假设 M1 为拥有 θ_1 个参数和对数似然函数 $L(\theta_1)$ 的未识别模型。再假设一个重新参数化后的模型 G 将已识别的维度和未识别的维度相分离。 $\theta_1 = G(\theta_2, \theta_3)$ 是一一对应可变的参数转化, 其中 θ_2 代表可以被识别的参数, θ_3 代表未识别的参数。 θ_1 代表模型中的观测指标截距和因子载荷, 以及潜因子均值, 潜因子方差, 因子间协方差。 θ_2 代表形态不变模型里的截距和因子载荷参数, θ_3 代表潜因子均值, 潜因子方差。

惩罚对齐法设定未识别的 θ_3 参数的所有潜因子均值为 0, 潜因子方差为 1, 可以得出一个可识别的零假设模型 M0, 也可以称作形态不变模型。形态不变模型有一些重要的性质, 比如模型 M0 和 M1 的对数似然值和模型拟合度是相等的, 自由估计的参数数量也是相等的。对于模型 M0, θ_3 参数被识别是因为潜因子均值被设定为定值 0, 潜因子方差被设定为定值 1。而对于模型 M1, θ_3 参数通过惩罚函数 $P(\theta)$ 被识别。惩罚对齐法分析的核心思想是使对数似然函数 $L(\theta)$ 取得最大值而 $P(\theta)$ 取得合理区间内的最小值, 完成测量不变性检验, 计算出模型 M1 中的潜因子均值和潜因子方差, 这样研究者就可以进行多组比较。

为了使加权的对数似然值取得最大值, 惩罚对齐法模型应满足下式:

$$L_w(\theta_1) = (1-w) L(\theta_1) - wP(\theta_1) \quad (7)$$

其中 w 代表惩罚函数的权重, 通常设为 0.1 或 0.01。一般来说, 惩罚函数的权重越小, 模型分析就需要更低的收敛标准。惩罚对齐法通过选择越来越小但不为 0 的惩罚函数权重 w , 使对数似

然函数 $L_w(\theta_1)$ 逐渐得到近似最优解。这里的 w 必须在合理区间, 不能取无限小的正值, 因为 w 过小可能导致模型无法收敛、无法识别、标准误估计出现鞍点等问题(Asparouhov & Muthén, 2023b)。

在惩罚对齐法模型中, 为了使加权的对数似然值 L_w 取得最大值, 加权的 $\theta_{2,w}$ 和 $\theta_{3,w}$ 参数需满足下式:

$$\frac{\partial L_w(\theta_2, \theta_3)}{\partial \theta_2} = (1-w) \frac{\partial L(\theta_2, \theta_3)}{\partial \theta_2} - w \frac{\partial P(\theta_2, \theta_3)}{\partial \theta_2} = 0 \quad (8)$$

通过此式只可以得出 $\theta_{2,w}$ 参数的解, 将此解记作 $\theta_{2,w}$, 此时求出了 $L(\theta_1)$ 的最大值。为了使 L_w 取得最大值, 对齐惩罚函数 $P(\theta)$ 应在合理范围内取得最小值。根据前文对于对齐法原理的描述, 利用形态不变模型 M0 和引入惩罚函数的模型 M1 具有相同的模型拟合度和对数似然值的性质, 依据对齐惩罚函数 $P(\theta)$ 选取相应的 $\theta_{3,w}$, 使公式(1)表示的总损失函数 F 取得最小值解。由于组样本量为已知确定值, 最小化总损失函数 F 本质上就是要让对齐惩罚函数 $P(\theta)$ 取得最小值。因此, $\theta_{3,w}$ 需满足下式:

$$\frac{\partial L_w(\theta_2, \theta_3)}{\partial \theta_3} = -w \frac{\partial P(\theta_2, \theta_3)}{\partial \theta_3} = 0 \quad (9)$$

当为参数设置均值为 μ , 先验分布方差为 v 的对齐先验分布 $ALF(\mu, v)$ 时, 对齐惩罚函数 $P(\theta)$ 如下式所示:

$$P(\theta) = \frac{f(\theta - \mu)}{v} \quad (10)$$

其中对齐成分损失函数 $f(x)$ 如下式所示:

$$f(x) = \sqrt[4]{x^2 + \varepsilon} \approx \sqrt{|x|} \quad (11)$$

在惩罚对齐法中需要将属于相同观测指标, 但属于不同组的因子载荷或截距两两间的差异纳入 $f(x)$ 进行优化计算。此处 ε 为一个很小的正值 (Mplus 软件中默认设为 0.001), 可以使对齐惩罚函数拥有连续一阶导数, 惩罚对数似然值的优化变得更简单。小的 ε 值意味着参数点估计值的偏差较小, 大的 ε 值意味着参数标准误估计值的偏差较小。如上文对齐法原理部分所述, 根据对齐成分损失函数 $f(x)$ 的性质, 公式(1)所示的总损失函数 F 取得最小值的解应当是当模型拥有少量较大不等参数和大多数满足近似测量不变性参数的时候。通过公式(9)至(11)可以得出 $\theta_{3,w}$ 参数的

解, 将此解记作 $\hat{\theta}_{3,w}$ 。M1 模型所有参数 $\theta_{1,w}$ 的解记作 $\theta_{1,w} = G(\theta_{2,w}, \theta_{3,w})$ 。

至此, 我们发现惩罚对齐法没有检验模型中的因子载荷和截距严格相等, 而是基于形态不变模型引入对齐惩罚函数 $P(\theta)$, 求得形态不变模型原本无法识别的潜因子均值和潜因子方差, 然后进一步选取最优测量不变模型完成测量不变性检验, 并进行多组比较。

3.2 惩罚对齐法继承传统方法的优势与克服的不足

用一句话来概括, 惩罚对齐法的主要优势是它继承并整合了多组探索性因子分析, 对齐法和贝叶斯结构方程模型的优势, 与此同时克服了这三种方法的不足。下面将具体介绍惩罚对齐法继承了传统方法的哪些优势, 克服了哪些不足。

3.2.1 继承了多组探索性因子分析的优势, 克服了其不足

多组探索性因子分析的典型方法是多组探索性结构方程模型。传统验证性因子分析忽略估计交叉载荷, 导致因子载荷和因子间协方差的高估, 潜因子均值估计出现偏差(Asparouhov & Muthén, 2009; Marsh et al., 2013; Wen & Hu, 2022), 探索性结构方程模型的提出解决了此问题(Marsh et al., 2013; Asparouhov & Muthén, 2023a)。一些研究使用目标旋转法(target rotation)将较小的不显著因子载荷设为近似为零的目标载荷, 取得了较好的分析效果。基于目标旋转法的多组探索性结构方程模型一度被认为是最好的多组探索性因子分析方法(Asparouhov & Muthén, 2023b)。惩罚对齐法继承了多组探索性因子分析在 EFA 框架内进行测量不变性检验和可以精确估计模型中的交叉载荷和近似为零因子载荷的优势。

同时, 多组探索性因子分析存在一些不足, 但惩罚对齐法克服了这些不足。第一, 使用多组探索性结构方程模型进行多组比较需要检验模型中的因子载荷和截距严格相等, 但现实中很难满足。而惩罚对齐法检验的是模型的近似测量不变性, 在多因子情境下允许模型中存在 20%至 30%的不等参数, 限制性更低。第二, 采用目标旋转法的多组探索性结构方程模型如果目标载荷出现误设, 模型参数估计效果会变得不好。而惩罚对齐法即使出现误设, 也能精确估计各类模型参数。第三, 采用目标旋转法的多组探索性结构方程模

型要求每个因子至少设定模型因子数 $m-1$ 个目标载荷才能使参数估计结果有较小的均方误差, 这就有可能出现被迫误设的情况。而惩罚对齐法可以直接对所有测量参数设定对齐先验分布, 不必满足最低目标载荷设置数量。

3.2.2 继承了对齐法的优势, 克服了其不足

对齐法基于形态不变模型, 通过检验模型参数的近似测量不变性达到潜因子均值比较的目的, 较好地解决了传统方法进行多组比较需要满足苛刻的因子载荷和截距严格相等假设的不足(温聪聪 等, 2019; Asparouhov & Muthén, 2014)。2023年, Asparouhov 和 Muthén 又开发出了对齐探索性结构方程模型, 该方法既允许多因子模型中存在 20% 至 30% 的不等参数, 又可以精确估计模型中的交叉载荷(Asparouhov & Muthén, 2023a)。将对齐法中的成分损失函数作为惩罚函数的惩罚对齐法当然也具备传统对齐法的优势。

同时, 对齐法存在一些不足, 但惩罚对齐法克服了这些不足。第一, 无论是传统对齐法还是对齐探索性结构方程, 在使用固定识别算法时需要将选择的参照组中所有潜因子的因子均值设定为 0, 潜因子方差设定为 1, 但现实中很难保证该组所有潜因子均值相等, 把它们都设为 0 就会出现误设, 影响模型参数估计效果。而惩罚对齐法不需要设置唯一的参照组, 即设为 0 的潜因子均值可以存在于不同组。第二, 对齐法无法应用于多指标多因素模型(MIMIC model), 即无法在模型中引入预测潜因子的观测指标, 但惩罚对齐法可以引入(Asparouhov & Muthén, 2023b)。

3.2.3 继承了贝叶斯结构方程模型的优势, 克服了其不足

传统多组比较方法进行测量不变性检验时很难满足因子载荷和截距严格相等的假设, 需要利用修正指数主观确定下一步分析释放哪些参数被自由估计, 模型可能越修正越远离真实的模型(Asparouhov & Muthén, 2014; Marsh et al., 2018)。在这一背景下, 贝叶斯结构方程模型被应用于测量不变性检验来解决上述问题。该方法一方面通过检验模型的近似测量不变性达到潜因子均值比较的目的, 放宽了对模型参数的限制。另一方面, 该方法可以对模型参数设定较小先验方差的信息先验分布, 将已有理论和预测提供的信息运用于分析中, 有效估计模型中的近似测量不变参数、

低程度交叉载荷和残差相关。惩罚对齐法在进行多组比较时也检验模型的近似测量不变性, 并且对模型参数设定对齐先验分布, 该方法同样具备贝叶斯结构方程模型的优势。

同时, 贝叶斯结构方程模型存在一些不足, 但惩罚对齐法克服了这些不足。第一, 贝叶斯结构方程模型是基于贝叶斯估计法, 而惩罚对齐法是基于极大似然估计法。在极大似然估计法框架中, 可以通过似然比检验比较嵌套模型拟合度的优劣, 而贝叶斯估计法框架不具备这一特点。同时, 同等模型极大似然估计法的计算速度也明显快于贝叶斯估计法。第二, 贝叶斯结构方程模型在进行测量不变性检验时检验的是近似测量不变性, 假设大部分跨组对应的模型参数近似相等, 所以对模型参数差异设置的是零均值正态先验分布。正态先验分布倾向于最小化待估计参数的均方误差, 可能把参数的不拟合性传导至整个模型, 影响参数估计效果(Asparouhov & Muthén, 2023b)。而惩罚对齐法对模型参数差异设置的是基于惩罚函数 $P(\theta) = \frac{f(\theta - \mu)}{v}$ 和对齐成分损失函

数 $f(x) \approx \sqrt{|x|}$ 的对齐先验分布, 对齐先验分布倾向于减少不等参数数量, 不假设先验分布满足正态分布, 也不太依赖先验方差的精细设定。由于对齐先验分布具有这些性质, 其应用于参数估计时更接近现实情况, 相比正态先验分布能提供更无偏的估计结果, 95%置信区间真值覆盖率通常也更高(Asparouhov & Muthén, 2023b)。第三, 贝叶斯结构方程模型在进行测量不变性检验时可以分为三步(Muthén & Asparouhov, 2013): (1)先对所有跨组对应模型参数之间的差异设定较小的先验分布方差, 例如 0.01 (宋琼雅 等, 2021; Kim et al., 2017), 得出 PPp、DIC、BIC 等模型拟合指标结果; (2)对所有跨组对应模型参数之间的差异设定较大的先验分布方差, 例如 0.05 或 0.1, 增大了参数差异的变异程度, 此时比较与设定先验分布方差为 0.01 模型的 PPp、DIC、BIC 等模型拟合指数是否有显著提升。如果先验方差为 0.01 的模型拟合指标更好或两个模型的拟合指标无显著差异, 则可以认为模型满足近似测量不变性; 否则, 说明模型不满足近似测量不变性。这前两个步骤通常被称为敏感性分析(宋琼雅 等, 2021; Muthén & Asparouhov, 2012); (3)根据敏感性分析选取出的

更优模型的估计结果对不满足近似测量不变性的少量参数差异调整为较大先验分布方差,得出修正后的最优近似测量不变性模型。而惩罚对齐法则是对所有跨组对应参数间的差异设置较大先验分布方差,一步就可以精确估计出所有的近似测量不变参数和不等参数。

通过将惩罚对齐法与传统测量不变性检验方法进行比较,可以看出惩罚对齐法具有明显优势。

4 研究实例

2014年5月至9月,北京大学“高等理科教育改革调研组”(《关于开展高等理科教育改革的调研的通知》高教司函[2012]221号)在全国百余所高校中进行了本科生网络问卷调查,并于9月初成功回收有效问卷100941份。本节将运用该问卷中大学生职业价值观量表数据的分析作为研究实例演示如何使用惩罚对齐法进行测量不变性检验,并比较985非C9高校、211高校、普通高校、C9高校大学生的职业价值观差异。测量大学生职业价值观的量表共设6题,使用4点Likert计分方式(1分为不重要、2分为不太重要、3分为比较重要、4分为重要)进行测量,设问为“考虑到您自己的未来发展,您认为下列因素的重要程度如何?”该量表包含两个因子,利他因子和自我实现因子各包含3个等级多分类观测指标,具体内容如表1所示。由于每个4点Likert计分变量在分析时包含3个阈值,一些填答类别的学生数过少,会导致模型无法识别。因此,将研究实例中的4点Likert计分变量转换为了3点计分等级多分类变量,1、2分记为“1”分,3、4分重新分别记为“2”分和“3”分。

表1 问卷中职业价值观量表的维度和题项设计

维度设计	题号	题项内容
职业价值观量表	Q1	有助于社会的可持续发展 and 环境保护
	Q2	帮助他人
	Q3	做对社会发展有重要意义的事情
	Q4	自我发展
	Q5	做我感兴趣的事情
	Q6	发挥我的天赋和能力

前文公式(1)表示的总损失函数 F 中的 $W_{g1,g2} = \sqrt{N_{g1}N_{g2}}$ 是用来反映组样本量大小和特定组参数估计的确定性,较大的组样本量比较小的组样本量对总损失函数的贡献更大。原始样本数据的总样本量为90416,其中985高校学生的样本量为16167,211高校学生的样本量为9249,普通本科高校学生的样本量为63585,C9高校学生的样本量为1415。根据 $W_{g1,g2} = \sqrt{N_{g1}N_{g2}}$ 的公式可知,在比较G3普通本科高校组和其他组参数的不等性时,G3对 $W_{g1,g3}$ 的贡献是G1的 $\sqrt{\frac{N_{g3}}{N_{g1}}} \approx 2.0$ 倍,

$W_{g2,g3}$ 的贡献是G2的 $\sqrt{\frac{N_{g3}}{N_{g2}}} \approx 2.6$ 倍,G3对 $W_{g3,g4}$

的贡献是G4的 $\sqrt{\frac{N_{g3}}{N_{g4}}} \approx 6.7$ 倍。为了防止对齐法

测量不等性优化计算中各组样本量差距过大而导致较大组对参数差异不等性的贡献过高,对原始样本中普通本科高校组和985高校组各重新随机抽样,随机选取保留9249个个案。这样一来,4个高校组的样本量比数变为6.5:6.5:6.5:1。985高校、211高校、普通高校组的参数在进行测量不等性优化计算时,它们之间的样本量权重是相等的。而这三个组在测量不等性优化计算中样本量影响的权重是C9院校组的 $\sqrt{\frac{N_{g3}}{N_{g4}}} = \sqrt{\frac{9249}{1415}} \approx 2.6$

倍。通过这样的重新随机抽样策略,平衡了各组样本量在测量不等性优化计算中的权重,明显缩小了各组样本量的差距。经过重新随机抽样,985非C9高校、211高校、普通高校、C9高校4种院校类型大学生的数量分别为9249,9249,9249,1415名。

4.1 探索性因子分析

首先对职业价值观量表进行探索性因子分析,检验研究模型的维度与问卷设计是否相符。使用Mplus软件运行探索性因子分析,得出表2中所示的旋转后因子载荷矩阵。题项1,2,3在 F_1 上有主载荷,题项4,5,6在 F_2 上有主载荷,旋转后的因子载荷矩阵体现的量表维度和问卷设计一致,可以判定本研究的职业价值观量表设计得到了样本数据的支持。

4.2 多组验证性因子分析

然后,使用Mplus软件运行多组验证性因子

表 2 旋转后的因子载荷矩阵(N = 29162)		
题项	F1	F2
Q1 有助于社会的可持续发展和环境保护	0.61	0.17
Q2 帮助他人	0.86	-0.01
Q3 做对社会发展有重要的事情	0.65	0.21
Q4 自我发展	0.03	0.84
Q5 做我感兴趣的事情	-0.09	0.93
Q6 发挥我的天赋和能力	0.16	0.70

分析进行测量不变性检验。对于使用 WLSMV 估计法进行的测量不变性检验，等级多分类数据的特殊性导致形态不变模型在不同识别方式下再继续增加因子载荷参数的测量不变限定会导致不等价的模型出现，使模型无法识别(Wu & Estabrook, 2016; Muthén & Muthén, 2023)。而对于等级多分类数据，在模型中为每个观测指标引入并解释连续型潜在反应变量，使潜在反应变量成为新的潜因子，将潜在反应变量的截距作为潜因子均值的载体，而潜在反应变量又可以解释原来的公因子维度。等级多分类数据进行潜因子均值比较的前提条件与连续变量数据相似，需要模型同时满足潜在反应变量因子载荷和观测指标阈值的测量不变性(Wu & Estabrook, 2016)。本例分别使用形态不变、阈值不变、阈值和因子载荷不变模型拟合数据，得到表 3 所示的拟合结果。可以看出，形态不变模型与阈值不变模型自由估计参数数目和自由度都相等，二者非嵌套模型，无法使用传统的卡方变化检验进行比较，而同时限定阈值和因子载荷不变的模型与形态不变模型和阈值不变模型

属于嵌套模型，可以使用 Mplus 软件中的 DIFFTEST 命令得出比较结果。根据卡方检验结果与 Svetina 和 Rutkowski 提出的在二因子、高组样本量、等级多分类指标情境下 $\Delta RMSEA$ 小于等于 0.05 即拒绝嵌套模型的判定标准(Svetina & Rutkowski, 2017)，竞争模型的 $\Delta RMSEA$ 为 0.015，不满足阈值和载荷不变性，无法直接比较各组的潜在反应变量截距。在传统多组验证性因子分析中，当组样本量较大时，卡方检验结果是较难不显著的。而 $\Delta RMSEA$ 、 ΔCFI 等判断嵌套模型是否被拒绝的拟合指数标准也有许多版本，很难用统一的标准去适配不同的研究情境。所以，竞争模型间的对比是传统测量不变性检验方法较难解决的问题。接下来，基于形态不变模型，尝试使用惩罚对齐法检验模型的近似测量不变性，以达成多组比较的目的。

4.3 基于验证性因子分析模型的惩罚对齐法分析

在 Mplus 8.10 版本软件中采用和 4.2 小节多组验证性因子分析部分相同的因子载荷矩阵结构，基于验证性因子分析模型运行惩罚对齐法(即忽略估计模型中所有的交叉载荷)，得出表 4 中的参数近似测量不变性分析结果。从模型拟合结果来看，RMSEA 为 0.065，CFI 为 0.993，TLI 为 0.988，SRMR 为 0.021，这些指标结果都很好。从近似测量不变性分析结果来看，除了 4 个组在 y_3 的第二个阈值 $\tau_{2,3}$ 都不满足近似测量不变性，大部分阈值参数都满足近似测量不变性，所有因子载荷都满足近似测量不变性，测量不变性检验的效果很好。根据这些结果，可以认为惩罚对齐法的分析效果较好，直接对各组的潜因子均值进行比较切实可行。

表 3 使用多组验证性因子分析进行测量不变性检验得到的模型拟合结果(N = 29162)								
模型	CFI	TLI	RMSEA/[90%CI]	SRMR	自由估计参数数目	卡方值	自由度	p 值
形态不变	0.993	0.988	0.077[0.073, 0.080]	0.021	76	1407.25	32	<0.001
阈值不变	0.993	0.988	0.077[0.073, 0.080]	0.021	76	1407.10	32	<0.001
阈值和载荷不变	0.994	0.992	0.062[0.059,0.064]	0.022	64	1257.38	44	<0.001
阈值不变对形态不变(变化值)	—	—	—	—	—	—	—	—
阈值和载荷不变对形态不变(变化值)	0.001	0.004	0.015	0.001	12	33.30	12	<0.001
阈值和载荷不变对阈值不变(变化值)	0.001	0.004	0.015	0.001	12	33.30	12	<0.001

注：表内的测量不变性检验模型均采用的是 theta 参数化分析。使用 WLSMV 估计法进行因子分析时可以选择 theta 参数化分析(非标准化的)或 delta 参数化分析(标准化的)，其区别是 theta 参数化分析出于模型识别目的将所有自由估计因子载荷和阈值的分类观测指标的残差方差设为定值 1，而 delta 参数化分析则是为每个自由估计因子载荷和阈值的分类观测指标设置一个为定值 1 的尺度因子。采用这两种参数化分析尺度的模型拟合度相同。

表4 模型参数的近似测量不变性分析结果($N = 29162$)

参数类型	满足与不满足近似测量不变性的组	参数类型	满足与不满足近似测量不变性的组
阈值		因子载荷	
$\tau_{1,1}$	1 2 3 4	$\lambda_{1,1}$	1 2 3 4
$\tau_{1,2}$	1 2 (3) (4)	$\lambda_{1,2}$	1 2 3 4
$\tau_{1,3}$	1 2 3 4	$\lambda_{1,3}$	1 2 3 4
$\tau_{1,4}$	1 2 3 4	$\lambda_{2,4}$	1 2 3 4
$\tau_{1,5}$	(1) 2 (3) 4	$\lambda_{2,5}$	1 2 3 4
$\tau_{1,6}$	(1) 2 (3) 4	$\lambda_{2,6}$	1 2 3 4
$\tau_{2,1}$	1 2 3 4		
$\tau_{2,2}$	1 2 3 4		
$\tau_{2,3}$	(1) (2) (3) (4)		
$\tau_{2,4}$	1 2 3 4		
$\tau_{2,5}$	(1) 2 (3) 4		
$\tau_{2,6}$	1 2 3 4		

注：数字加括号标粗体表示该组的该因子载荷或阈值参数不满足近似测量不变性。

在本例中，假设最初的形态不变模型是未识别的模型 M1，它包含 θ_1 个参数，其中 $\theta_1 = \theta_2 + \theta_3$ ， θ_2 代表可以被识别的因子载荷和阈值参数， θ_3 代表未识别的潜因子均值和潜因子方差参数。多组验证性因子分析中的形态不变模型由于要估计因子载荷和阈值，通常设置潜因子均值为定值 0，潜因子方差为定值 1，此时的形态不变模型为 M0。使用惩罚对齐法的对齐惩罚函数 $P(\theta)$ 在 M1 模型中分别对因子载荷参数和截距参数的两两跨组差异设置对齐先验分布，利用 M0 和 M1 模型有相等似然值的性质使未识别的维度(θ_3 参数，在本例中即潜因子均值，潜因子方差)变得可识别，得出的对齐模型便是可识别的 M1 模型。对齐惩罚函数的特点是追求简洁的近似测量不变性结构，使少数模型参数有较大不等性，大多数模型参数近似测量不变，这一特点通常符合量表设计。

在 Mplus 软件中设置新参数进行双样本均值 Z 检验计算，并且为避免多次两两比较时出现一类错误率膨胀问题，采用 Bonferroni 方法调整每两组比较时使用的显著性水平 α 。4 个组在每个潜因子进行两两比较需比较 $C_4^2 = 6$ 次，Bonferroni 方法调整后的显著性水平 $\alpha = 0.05/6 \approx 0.008$ 。使用调整后的显著性水平进行两两比较，总结出表 5。可以看出，C9 高校学生的自我实现潜因子均值

最高，利他潜因子均值最低；普通高校学生的利他潜因子均值高于其他院校类型的学生，但自我实现潜因子均值较低；211 高校学生的自我实现潜因子均值最低，利他潜因子均值也较低，说明样本中该院校类型的学生对这两项职业价值观的看重程度总体是比较消极的态度。

表5 不同院校类型学生职业价值观潜因子均值的比较结果($N = 29162$)

所属因子	排名	组名称	因子均值	因子均值显著小于该组的组名称
利他因子	1	普通高校	0.05	211 C9
	2	985	0	211 C9
	3	211	-0.08	
	4	C9	-0.18	
自我实现因子	1	C9	0.04	普通高校 211
	2	985	0	普通高校 211
	3	普通高校	-0.13	
	4	211	-0.19	

4.4 基于探索性结构方程模型的惩罚对齐法分析

在 Mplus 8.10 版本软件中基于探索性结构方程模型运行惩罚对齐法(即为模型中的因子载荷设置 EFA 旋转函数，并且估计模型中所有的交叉载荷)，得出表 6 中的参数近似测量不变性分析结

表6 模型参数的近似测量不变性分析结果($N = 29162$)

参数类型	满足与不满足近似测量不变性的组	参数类型	满足与不满足近似测量不变性的组
阈值		因子载荷	
$\tau_{1,1}$	1 2 3 4	$\lambda_{1,1}$	(1) 2 3 4
$\tau_{1,2}$	(1) 2 3 4	$\lambda_{1,2}$	1 2 3 4
$\tau_{1,3}$	(1) 2 (3) 4	$\lambda_{1,3}$	(1) 2 3 4
$\tau_{1,4}$	1 2 3 4	$\lambda_{1,4}$	1 (2) 3 4
$\tau_{1,5}$	1 (2) 3 4	$\lambda_{1,5}$	1 2 3 4
$\tau_{1,6}$	1 (2) 3 (4)	$\lambda_{1,6}$	1 2 3 4
$\tau_{2,1}$	1 2 3 4	$\lambda_{2,1}$	1 (2) 3 4
$\tau_{2,2}$	1 2 3 4	$\lambda_{2,2}$	1 (2) 3 4
$\tau_{2,3}$	(1) 2 (3) 4	$\lambda_{2,3}$	1 2 3 4
$\tau_{2,4}$	1 2 3 4	$\lambda_{2,4}$	1 2 3 4
$\tau_{2,5}$	1 2 3 4	$\lambda_{2,5}$	1 (2) 3 4
$\tau_{2,6}$	1 2 3 4	$\lambda_{2,6}$	(1) 2 3 4

注：数字加括号标粗体表示该组的该因子载荷或阈值参数不满足近似测量不变性。

果。从模型拟合结果来看, RMSEA 为 0.067, CFI 为 0.994, TLI 为 0.978, SRMR 为 0.018, 这些指标结果都很好。从近似测量不变性分析结果来看, 大部分阈值参数和因子载荷都满足近似测量不变性, 测量不变性检验的效果较好。根据这些结果, 可以认为惩罚对齐法的分析效果较好, 直接对各组的潜因子均值进行比较切实可行。

在 Mplus 软件中设置新参数进行双样本均值 Z 检验计算, 使用 Bonferroni 方法调整后的显著性水平进行两两比较, 总结出表 7。可以看出, C9 高校学生的自我实现潜因子均值最高, 利他潜因子均值最低; 普通高校学生的利他潜因子均值高于其他院校类型的学生, 但自我实现潜因子均值较低; 211 高校学生的自我实现潜因子均值最低, 利他潜因子均值也较低, 说明样本中该院校类型的学生对这两项职业价值观的看重程度总体是比较消极的态度。

表 7 不同院校类型学生职业价值观潜因子均值的比较结果(N = 29162)

所属因子	排名	组名称	因子均值	因子均值显著小于该组的组名称
利他因子	1	普通高校	0.08	985 211 C9
	2	985	0	C9
	3	211	-0.07	C9
	4	C9	-0.27	
自我实现因子	1	C9	0.15	985 普通高校 211
	2	985	0	普通高校 211
	3	普通高校	-0.18	
	4	211	-0.21	

4.5 不同模型分析结果的比较

本节使用了多组验证性因子分析、基于验证性因子分析模型的惩罚对齐法分析和基于探索性结构方程模型的惩罚对齐法分析分别拟合了样本数据, 取得了各自的分析结果。接下来, 需要比较这三种模型的分析结果。

首先, 基于验证性因子分析模型的惩罚对齐法分析结果优于多组验证性因子分析的结果。第一, 基于验证性因子分析模型的惩罚对齐法分析的模型参数近似测量不变性结果显示所有的因子载荷都满足近似测量不变性, 意味着模型满足弱测量不变性。而传统的多组验证性因子分析使用

WLSMV 估计法检验的因子载荷和截距严格相等的模型被拒绝, 且因子载荷严格相等的模型无法识别(Wu & Estabrook, 2016; Muthén & Muthén, 2023), 这就导致传统多组验证性因子分析得出的模型和数据样本的实际情况相差甚远。第二, 基于验证性因子分析模型的惩罚对齐法分析不仅得出了所有模型参数是否满足近似测量不变性的诊断结果, 并且可以基于形态不变模型估计出模型中的潜因子均值参数进行直接比较。而传统多组验证性因子分析进行的嵌套模型的卡方似然比检验表明因子载荷和截距严格相等的假设被拒绝, 无法进行直接的潜因子均值比较。结合惩罚对齐法分析提供的参数近似测量不变性结果来看, 多组验证性因子分析使用修正指数从后往前不断修正, 得出部分因子载荷和截距严格相等模型是个复杂的过程, 并且可能由于最初模型的误设、修正参数的主观选择和修正指数的多重共线性导致修正过程并不一定正确。

其次, 基于验证性因子分析模型和基于探索性结构方程模型的惩罚对齐法分析结果总体一致, 但基于探索性结构方程模型的惩罚对齐法分析结果更能反映数据样本的真实情况。第一, 4 个院校类型组在两个潜因子上的潜因子均值排序完全一致, 但各组潜因子均值的绝对数值差异有所不同, Z 检验两两比较的显著性结果略有不同。例如, 基于验证性因子分析模型的惩罚对齐法分析结果显示 985 高校学生的利他潜因子均值显著高于 211 高校学生, 但在基于探索性结构方程模型的惩罚对齐法分析中这一差异不显著。基于探索性结构方程模型的惩罚对齐法分析结果显示普通高校学生的利他潜因子均值显著高于 985 高校学生, 211 高校学生的利他潜因子均值显著高于 C9 高校学生, C9 高校学生的自我实现潜因子均值显著高于 985 高校学生, 但在基于验证性因子分析模型的惩罚对齐法分析中这些差异不显著。第二, 从模型拟合结果来看, 基于探索性结构方程模型的惩罚对齐法分析的 CFI, SRMR 拟合指标表现更好, 基于验证性因子分析模型的惩罚对齐法分析的 TLI, RMSEA 拟合指标表现更好。由于两个模型并非嵌套模型, 从卡方值来看, 基于验证性因子分析模型的惩罚对齐法分析的卡方值为 1403.827, 自由度为 32, 基于探索性结构方程模型的惩罚对齐法分析的卡方值为 1219.664, 自由度为 16, 两

个模型卡方差值较大,说明基于探索性结构方程模型的惩罚对齐法分析的效果可能更好。第三,从近似测量不变性检验的结果来看,基于验证性因子分析模型的惩罚对齐法分析的不等参数率为 $12/72 \approx 16.7\%$,基于探索性结构方程模型的惩罚对齐法分析的不等参数率为 $13/96 \approx 13.5\%$,后者的近似测量不变性更高。不过,前者的所有因子载荷全都满足近似测量不变性,并且该模型较为简洁,估计的自由参数较少,读者可能会觉得前者的近似测量不变性检验结果更好。实际上,验证性因子分析模型本身忽略估计了交叉载荷,而基于探索性结构方程模型的惩罚对齐法分析的近似测量不变性检验结果表明这些忽略估计的交叉载荷中有3个不满足近似测量不变性,基于探索性结构方程模型估计交叉载荷才能更好地反映真实数据情况。

综上所述,对于本研究实例中的大学生职业价值观量表数据,采用基于探索性结构方程模型的惩罚对齐法分析数据应是更好的选择。

5 总结与讨论

惩罚结构方程模型是一种全新的模型框架,能应用于许多不同的情境。将其应用于测量不变性检验和多组比较的典型案例是惩罚对齐法。在惩罚对齐法提出之前,贝叶斯结构方程模型、对齐法和多组探索性结构方程等模型是研究者们用于测量不变性检验和多组比较的常用方法,而惩罚对齐法的提出对这些传统方法造成了冲击。

本研究首先回顾了几种常用的测量不变性检验方法的原理、优势和不足,其次引入惩罚结构方程模型框架和惩罚对齐法,详述了惩罚对齐法的运算过程,总结了惩罚对齐法从多组探索性因子分析、对齐法、贝叶斯结构方程模型等传统方法继承的优势,克服了这些方法的哪些不足,最后运用大学生职业价值观研究实例比较了多组验证性因子分析、基于验证性因子分析的惩罚对齐法分析、基于探索性结构方程模型的惩罚对齐法分析拟合样本数据的效果,演示如何使用惩罚对齐法进行测量不变性检验和多组比较。

概括来说,惩罚对齐法继承了多组探索性因子分析在EFA框架内进行测量不变性检验和可以估计交叉载荷等优点,继承了对齐法使用成分损失函数,允许模型中存在一定量不等参数等优点,

继承了贝叶斯结构方程模型对模型参数设置先验分布,检验模型参数的近似测量不变性等优点。同时,该方法克服了多组探索性因子分析需要满足因子载荷和截距严格相等的假设、使用目标旋转法时容错率低的不足,克服了对齐法在多因子模型情境下设置潜因子均值参照组时容易出现误设、无法应用于多指标多因素模型的不足,克服了在使用贝叶斯结构方程模型进行测量不变性检验情境下设定的参数差异先验分布需满足正态分布假设、需要对模型进行敏感性分析、运算较慢等不足。由于惩罚对齐法整合了这些特点,其相比传统测量不变性检验方法具有明显的优势,在未来将有广阔的应用前景。

与此同时,惩罚对齐法也存在一些不足之处。例如,正是由于惩罚对齐法继承了多组探索性因子分析、对齐法、贝叶斯结构方程模型等测量不变性检验方法的优势,所以研究者在分析数据前需要对传统测量不变性检验方法比较熟悉,才能对惩罚对齐法的理解和使用有更好的把握。其次,基于多水平分析的二水平随机效应法具有一些独特的优势,但惩罚对齐法目前尚未与多水平分析模型进行整合,这也是该方法未来发展的方向之一。

惩罚对齐法的使用和传统对齐法较为相似,主要区别是惩罚对齐法可以在模型中引入交叉载荷,对参数或参数差异设置对齐先验分布。目前,该方法已经可以在最新的Mplus 8.10版本中使用。本研究演示的研究实例和附录提供的语法示例将能为研究者和读者提供参考,更方便地使用该方法进行测量不变性检验和多组比较。

参考文献

- 白新文,陈毅文.(2004). 测量等价性的概念及其判定条件. *心理科学进展*, 12(2), 231-239.
- 刘军,吴维库.(2005). 心理测量平衡性研究与实例. *心理科学*, 28(1), 170-174.
- 宋琼雅,张沥今,潘俊豪.(2021). 贝叶斯多组比较——渐近测量不变性. *心理学探新*, 41(1), 69-75.
- 王孟成.(2014). *潜变量建模与Mplus应用*. 重庆大学出版社.
- 王孟成,邓俏文,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.
- 王阳,温忠麟,付媛姝.(2020). 等效性检验——结构方程模型评价和测量不变性分析的新视角. *心理科学进展*, 28(11), 1961-1969.
- 王阳,温忠麟,李伟,方杰.(2022). 新世纪20年国内结构方程模型方法研究与模型发展. *心理科学进展*, 30(8),

- 1715–1733.
- 温聪聪, 史秋衡. (2022). 对齐法在单因子和三因子多组比较中允许的不等参数率. *心理科学*, 45(5), 1230–1242.
- 温聪聪, 伍伟平, 林光杰. (2019). 对齐(Alignment)——一种新的多群组分析法. *心理科学进展*, 27(1), 181–189.
- 张沥今, 陆嘉琦, 魏夏琰, 潘俊豪. (2019). 贝叶斯结构方程模型及其研究现状. *心理科学进展*, 27(11), 1812–1825.
- Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397–438.
- Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(4), 495–508.
- Asparouhov, T., & Muthén, B. (2023a). Multiple group alignment for exploratory and structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 30(2), 169–191.
- Asparouhov, T., & Muthén, B. (2023b). *Penalized structural equation models*. Technical Report. 2023. Retrieved Jan 2, 2024, from <http://www.statmodel.com/download/PML.pdf>
- Byrne, B. M., & Van de Vijver, F. J. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, 10(2), 107–132.
- Caycho-Rodríguez, T., Vilca, L. W., Valencia, P. D., Carbajal-Leon, C., Vivanco-Vidal, A., Saroli-Aranibar, D., ... Gallegos, W. L. A.. (2021). Cross-cultural validation of a new version in Spanish of four items of the preventive COVID-19 infection behaviors scale (PCIBS) in twelve Latin American countries. *Frontiers in Psychology*, 12, 763993, 1–20.
- Deng, L., & Yuan, K.-H. (2016). Comparing latent means without mean structure models: A projection-based approach. *Psychometrika*, 81, 802–829.
- 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.
- Jang, S., Kim, E. S., Cao, C., Allen, T. D., Cooper, C. L., Lapierre, L. M., ... Woo, J. -M. (2017). Measurement invariance of the satisfaction with life scale across 26 countries. *Journal of Cross-Cultural Psychology*, 48(4), 560–576.
- 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.
- Lomazzi, V. (2018). Using alignment optimization to test the measurement invariance of gender role attitudes in 59 countries. *Methods, Data, Analyses*, 12(1), 77–103.
- Marsh, H. W., Guo, J., Parker, P. D., Nagengast, B., Asparouhov, T., Muthén, B., & Dicke, T. (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.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Morin, A. J., & Von Davier, M. (2013). Why item parcels are (almost) never appropriate: Two wrongs do not make a right—Camouflaging misspecification with item parcels in CFA models. *Psychological Methods*, 18(3), 257–284.
- Millsap, R. E. (2012). *Statistical approaches to measurement invariance*. Routledge.
- Munck, I., Barber, C., & Torney-Purta, J. (2018). Measurement invariance in comparing attitudes toward immigrants among youth across Europe in 1999 and 2009: The alignment method applied to IEA CIVED and ICCS. *Sociological Methods & Research*, 47(4), 687–728.
- Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313–335.
- Muthén, B., & Asparouhov, T. (2013). *BSEM measurement invariance analysis*. Retrieved Jan 2, 2024, from <https://www.statmodel.com/examples/webnotes/webnote17.pdf>
- Muthén, B., & Asparouhov, T. (2018). Recent methods for the study of measurement invariance with many groups: Alignment and random effects. *Sociological Methods & Research*, 47(4), 637–664.
- Muthén, L. K., & Muthén, B. O. (2023). *Version 8.9 and 8.10 Mplus language addendum*. Los Angeles, CA: Author. Retrieved Jan 2, 2024, from <https://www.statmodel.com/download/Version%208.9%20and%208.10%20Addendum>
- Shi, D., Song, H., DiStefano, C., Maydeu-Olivares, A., McDaniel, H. L., & Jiang, Z. (2019). Evaluating factorial invariance: An interval estimation approach using Bayesian structural equation modeling. *Multivariate Behavioral Research*, 54(2), 224–245.
- Svetina, D., & Rutkowski, L. (2017). Multidimensional measurement invariance in an international context: Fit measure performance with many groups. *Journal of Cross-Cultural Psychology*, 48(7), 991–1008.
- Wen, C., & Hu, F. (2022). Investigating the Applicability of Alignment—A Monte Carlo Simulation Study. *Frontiers in Psychology*, 13, 845721, 1–13.
- Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*, 81(4), 1014–1045.
- Yuan, K.-H., Chan, W., Marcoulides, G. A., & Bentler, P. M. (2016). Assessing structural equation models by equivalence testing with adjusted fit indexes. *Structural Equation Modeling: A Multidisciplinary Journal*, 23(3), 319–330.

A new measurement invariance test method: Penalized alignment

WEN Congcong

(International College, Xiamen University, Xiamen 361005, China)

Abstract: In 2023, Asparouhov and Muthén proposed a new framework for structural equation modeling called the penalized structural equation modeling (PSEM). The penalized alignment method exemplifies the utilization of PSEM within the field of measurement invariance testing. The penalized alignment method inherits the advantage of estimating cross-loadings from multiple-group exploratory factor analysis. It inherits the advantages of alignment method using alignment component loss function, and allowing for the existence of 20% to 30% noninvariant parameters in the model. It also incorporates the advantages of Bayesian structural equation modeling in setting prior distributions for model parameters and testing approximate measurement invariance of model parameters. Additionally, the penalized alignment method addresses limitations present in traditional measurement invariance testing methods. To illustrate the application of penalized alignment method, a study on work values among college students is employed as an example to showcase the utilization of the penalized alignment method for conducting measurement invariance testing and multiple-group analysis. Multiple-group CFA, CFA based penalized alignment, and ESEM based penalized alignment models are used to fit the data.

Keywords: measurement invariance testing, penalized alignment, Bayesian structural equation modeling, exploratory structural equation model, alignment loss function prior

附录

附录 1：探索性因子分析(EFA)的 Mplus 语法

```
data:
FILE = C:\Users\wenco\Desktop\PSEM\3dele9249.csv; !指示数据路径
variable:
names are g y1-y6;          !命名数据中的变量
categorical=y1-y6;          !说明 y1-y6 是分类变量
usevariables are y1-y6;      !说明需要使用的变量
ANALYSIS:
type = EFA 1 3;              !指示 EFA 潜因子数分别为 1,2,3 进行分析
estimator = wlsmv;           !指示参数估计法为修正均值与方差加权最小二乘法
output: tech1;
```

附录 2：形态不变模型(多组 CFA) 的 Mplus 语法

```
DATA:
FILE = C:\Users\wenco\Desktop\PSEM\3dele9249.csv; !指示数据路径
VARIABLE:
NAMES = g y1-y6;              !命名数据中的变量
categorical=y1-y6;            !说明 y1-y6 是分类变量
USEVARIABLES = g y1-y6;       !说明需要使用的变量
grouping = g (1=g1 2=g2 3=g3 4=g4); !命名已知组
ANALYSIS:
estimator = wlsmv;            !指示参数估计法为修正均值与方差加权最小二乘法
PARAMETERIZATION = theta;     !指示参数运算方法
MODEL:
!model 命令下设定适用于所有组的总模型
v1 by y1@1; v2 by y2@1; v3 by y3@1; v4 by y4@1; v5 by y5@1; v6 by y6@1;
!设定潜在反应变量，多分类观测指标通过阈值使潜在反应变量取得离散值
f1 by v1-v3*; f2 by v4-v6*;
!置于参数右侧表示该参数被自由估计
[v1-v6@0];
!设定潜在反应变量的截距，第一组设为 0，其他组自由估计
v1-v6@0;
[y1$1-y6$1]; [y1$2-y6$2];
f1-f2@1;
!由于潜因子没有设为定值 1 的参照因子载荷，潜因子方差需设为 1
[f1-f2@0];
y1-y6@1;
!theta 参数运算方法要求第一组的残差方差设为定值 1
f1 with f2;
MODEL g2:
! model+组编号命令下设定该组的模型，该组命令与总模型命令冲突时以该组命令为准
f1 by v1-v3*; f2 by v4-v6*;
[v1-v6@0];
v1-v6@0;
[y1$1-y6$1]; [y1$2-y6$2];
```

```

f1-f2@1;
[f1-f2@0];
y1-y6@1;
!theta 参数化分析出于模型识别目的将所有自由估计因子载荷
和阈值的分类观测指标的残差方差设为定值 1

f1 with f2;
MODEL g3:
同 g2
MODEL g4:
同 g2
OUTPUT: tech1;sampstat;
savedata:DIFFTEST=deriv.csv; !指示保存形态不变模型 H1 中的导数, 矩阵, 拟合优度函数值,
                                自由度, 样本统计量数, 组数等信息

```

附录 3: 阈值不变模型(多组 CFA) 的 Mplus 语法

```

DATA:
同附录 2
VARIABLE:
同附录 2
ANALYSIS:
同附录 2
DIFFTEST=deriv.csv; !指示利用阈值不变模型 H1 的信息和附录 2 形态不变模型 H0 保存
                    的信息计算出修正的均值和方差, 进行 WLSMV 估计法情境下的卡
                    方变化检验

MODEL:
v1 by y1@1;
..... !设置潜在反应变量, 同附录 2
f1 by v1-v3*; f2 by v4-v6*;
[v1-v6@0];
v1-v6@0;
[y1$1-y6$1] (tau1-tau6); [y1$2-y6$2] (tau7-tau12); !设定阈值, 小括号中为阈值贴标签, 方便设置阈
                                                    值不变性

f1-f2@1;
[f1-f2@0];
y1-y6@1;
f1 with f2;
MODEL g2:
f1 by v1-v3*; f2 by v4-v6*;
[v1-v6*]; !设定潜在反应变量的截距
v1-v6@0;
[y1$1-y6$1](tau1-tau6); [y1$2-y6$2](tau7-tau12);
f1-f2@1;
[f1-f2@0];
y1-y6; !特定组的阈值和因子载荷都自由估计时残差方差设为定值 1, 特

```

定组的阈值或因子载荷设置跨组相等时残差方差自由估计。本例中设置了阈值跨组相等

```
f1 with f2;
MODEL g3:
同 g2
MODEL g4:
同 g2
OUTPUT:
tech1;sampstat;
savedata:DIFFTEST=deriv2.csv; !指示保存阈值不变模型 H1 中的导数, 矩阵, 拟合优度函数值, 自由度, 样本统计量数, 组数等信息
```

附录 4: 阈值和载荷不变模型(多组 CFA) 的 Mplus 语法

```
DATA:
同附录 2
VARIABLE:
同附录 2
ANALYSIS:
同附录 2
DIFFTEST=deriv2.csv; !指示利用阈值和载荷不变模型 H1 的信息和附录 3 阈值不变模型 H0 保存的信息计算出修正的均值和方差, 进行 WLSMV 估计法情境下的卡方变化检验

MODEL:
v1 by y1@1;
..... !设置潜在反应变量, 同附录 2
f1 by v1-v3*(lam1-lam3); f2 by v4-v6*(lam4-lam6);
[v1-v6@0];
v1-v6@0;
[y1$1-y6$1*](tau1-tau6); [y1$2-y6$2*](tau7-tau12);
f1-f2@1;
[f1-f2@0];
y1-y6@1;
f1 with f2;
MODEL g2:
f1 by v1-v3*(lam1-lam3); f2 by v4-v6*(lam4-lam6);
[v1-v6*];
v1-v6@0;
[y1$1-y6$1*](tau1-tau6); [y1$2-y6$2*](tau7-tau12);
f1-f2;
[f1-f2@0];
y1-y6;
f1 with f2;
MODEL g3:
```

附录 5: 基于验证性因子分析模型的惩罚对齐法 Mplus 语法

data:	
同附录 1	
variable:	
同附录 1	
ANALYSIS:	
estimator = wlsmv;	!指示参数估计法为修正均值与方差加权最小二乘法
iterations =10000;	!指示运算迭代数量
param=theta;	!指示参数运算方法
convergence=0.000005;	!设定模型收敛标准
tolerance=0.01;	!设定对齐法惩罚函数的容忍值
starts=200;	!设定随机初始值数量
model:	!model 命令下设定适用于所有组的总模型
f1 by y1-y3*1; f2 by y4-y6*1;	
f1 with f2;	
y1-y6@1;	!theta 参数运算方法要求因子载荷和阈值同时自由估计时观测指标残差方差设为定值 1
MODEL G1:	!model+组编号命令下设定该组的模型，该组命令与总模型命令冲突时以该组命令为准
f1 by y1-y3(lam1_1-lam1_3); f2 by y4-y6(lam1_4-lam1_6);	!括号内设定因子载荷参数标签
[y1\$1-y6\$1](tau1_1-tau1_6); [y1\$2-y6\$2](tau1_7-tau1_12);	!括号内设定阈值参数标签
[f1-f2@0];	
f1-f2@1;	
f1 with f2;	
MODEL G2:	
f1 by y1-y3(lam2_1-lam2_3); f2 by y4-y6(lam2_4-lam2_6);	
[y1\$1-y6\$1](tau2_1-tau2_6); [y1\$2-y6\$2](tau2_7-tau2_12);	
[f1-f2](m2 n2);	
f1-f2;	
f1 with f2;	
MODEL G3:	
f1 by y1-y3(lam3_1-lam3_3); f2 by y4-y6(lam3_4-lam3_6);	
[y1\$1-y6\$1](tau3_1-tau3_6); [y1\$2-y6\$2](tau3_7-tau3_12);	

```

[f1-f2](m3 n3);
f1-f2;
f1 with f2;
MODEL G4:
f1 by y1-y3(lam4_1-lam4_3); f2 by y4-y6(lam4_4-lam4_6);
[y1$1-y6$1](tau4_1-tau4_6); [y1$2-y6$2](tau4_7-tau4_12);
[f1-f2](m4 n4);
f1-f2;
f1 with f2;
MODEL PRIOR:
DO(1,3) DIFF(lam1_#-lam4_#)~ALF(0,2);
DO(4,6) DIFF(lam1_#-lam4_#)~ALF(0,2);
DO(1,6) DIFF(tau1_#-tau4_#)~ALF(0,2); DO(7,12) DIFF(tau1_#-tau4_#)~ALF(0,2);
MODEL CONSTRAINT:
NEW(DIFM12 DIFM13 DIFM14 DIFM23 DIFM24 DIFM34 DIFN12 DIFN13
DIFN14 DIFN23 DIFN24 DIFN34);
DIFM12=-m2; DIFM13=-m3; DIFM14=-m4;
DIFM23=m2-m3; DIFM24=m2-m4; DIFM34=m3-m4;
DIFN12=-n2; DIFN13=-n3; DIFN14=-n4;
DIFN23=n2-n3; DIFN24=n2-n4; DIFN34=n3-n4;
output: align;svalues;

```

!先验分布命令
!Do 命令说明是要分析第 1 至第 3 个参数, DIFF 命令说明是要分析参数间的差异, ALF 命令说明设置的先验分布是对齐先验分布
!使用 NEW 命令为模型中的潜因子均值差异命名
!对潜因子均值差异进行具体定义
!align 命令要求输出基于对齐先验分布的参数差异估计结果和近似测量不变性结果, svalues 命令要求保存参数估计结果

附录 6: 基于探索性结构方程模型的惩罚对齐法 Mplus 语法

```

data:
同附录 1
variable:
同附录 1
ANALYSIS:
同附录 5
model:
f1 by y1-y6; f2 by y1-y6;
f1 with f2;
y1-y6@1;
MODEL G1:

```

!model 命令下设定适用于所有组的总模型
!theta 参数运算方法要求因子载荷和阈值同时自由估计时观测指标残差方差设为定值 1
!model+组编号命令下设定该组的模型, 该组命令与总模型命令冲突时以该组命令为准

```

f1 by y1-y6(lam1_1-lam1_6); f2 by y1-y6(lam1_7-lam1_12); !括号内设定因子载荷参数标签
[y1$1-y6$1](tau1_1-tau1_6); [y1$2-y6$2](tau1_7-tau1_12); !括号内设定阈值参数标签
[f1-f2@0];
f1-f2@1;
f1 with f2;
MODEL G2:
f1 by y1-y6(lam2_1-lam2_6); f2 by y1-y6(lam2_7-lam2_12);
[y1$1-y6$1](tau2_1-tau2_6); [y1$2-y6$2](tau2_7-tau2_12);
[f1-f2](m2 n2);
f1 with f2;
f1-f2;
MODEL G3:
f1 by y1-y6(lam3_1-lam3_6); f2 by y1-y6(lam3_7-lam3_12);
[y1$1-y6$1](tau3_1-tau3_6); [y1$2-y6$2](tau3_7-tau3_12);
[f1-f2](m3 n3);
f1 with f2;
f1-f2;
MODEL G4:
f1 by y1-y6(lam4_1-lam4_6); f2 by y1-y6(lam4_7-lam4_12);
[y1$1-y6$1](tau4_1-tau4_6); [y1$2-y6$2](tau4_7-tau4_12);
[f1-f2](m4 n4);
f1 with f2;
f1-f2;
MODEL PRIOR:
DO(1,6) DIFF(lam1_#-lam4_#)~ALF(0,0.5);
DO(7,12) DIFF(lam1_#-lam4_#)~ALF(0,0.5);
DO(1,12) DIFF(tau1_#-tau4_#)~ALF(0,0.5);
lam1_1-lam1_12~Geomin(2,0.5,.001); lam2_1-lam2_12~Geomin(2,0.5,.001);
lam3_1-lam3_12~Geomin(2,0.5,.001); lam4_1-lam4_12~Geomin(2,0.5,.001);
MODEL CONSTRAINT:
同附录 5
output: align;svalues;

```

!先验分布命令

!Do 命令说明是要分析第 1 至第 6 个参数, DIFF 命令说明是要分析参数间的差异, ALF 命令说明设置的先验分布是对齐先验分布

!通过 Geomin 旋转函数为模型中的因子载荷参数设定使用探索性结构方程模型进行分析

!align 命令要求输出基于对齐先验分布的参数差异估计结果和近似测量不变性结果, svalues 命令要求保存参数估计结果