第 21 卷第 5 期　　经 济 学（季 刊）　　Vol. 21, No. 5
2021 年 9 月　　China Economic Quarterly　　September, 2021

调节效应的陷阱

朱家祥　张文睿*

摘　要　实证研究中，常在线性回归中加入交叉项来对调节效应建模。然而基于交叉项系数的统计显著来宣称调节效应，存在可能的误导。调节效应不能通过纯粹的统计分析来证实。如果缺乏可信的理论，以数据挖掘的方式来搜索调节效应很容易掉入虚假调节效应的陷阱。本文提出一系列预检验来避免虚假的调节效应，并建议了研究调节效应的实证策略。

关键词　交叉项回归，随机系数模型，调节效应
DOI：10.13821/j.cnki.ceq.2021.05.19

一、前　　言

交叉项的解释变量经常被用来捕捉 x 对 y 的影响决定于另一变量 z，这个变量 z 被称为调节变量（moderator）。此外，本文也用“被调节变量”来称呼解释变量 x 。考虑调节效应（moderating effect）的实证在社会科学的领域，例如心理学和市场学，有着广泛的探讨。经济金融的实证文献也常涉及调节效应的探讨。对调节效应建模的主要手段就是在线性回归引入交叉相乘项，并以交叉相乘项的回归系数是否显著来主张或拒绝调节效应。

我们可以用下列几个问题来阐述这个实证策略的任意性。假设起始模型包含了两个解释变量 x 和 w，同时令 z 为考虑中的调节变量。问题是：(1) 调节效应是基于理论或仅仅是改进模型的数据挖掘？(2) 交叉相乘项是谁和 z 交叉相乘，x 或是 w？或甚至是应该有两个交叉相乘项，即 xz 和 wz？(3) 调节变量 z 是如何选取的？只有一个，或是有多个候选的调节变量？这些问题在本文的后续会有详尽的讨论。

如果调节效应有一般认知或理论的支持，例如 Castro *et al*. (2004)，问题 (1) 命令了研究者必须在起始模型就引入交叉项作为解释变量。如果缺乏调节效应的理论，或者说缺乏具有说服力的观点，引入交叉项属于数据挖掘。

* 朱家祥、张文睿，北京大学汇丰商学院。通信作者及地址：朱家祥，广东省深圳市南山区深圳大学城北京大学汇丰商学院 732 室，518055；电话：(0755) 26032933；E-mail：cchu@phbs.pku.edu.cn。

如果是数据挖掘但却无法回答上述(2)和(3)的问题,即用于挖掘调节效应的模型没有容纳多个可能的调节变量及多个可能的被调节变量的一般性,那么数据挖掘可能挖得还不够深(线性回归甚至并不一定是数据挖掘的合适工具)。简言之,许多实证研究,说理论是既没有理论,说数据挖掘也挖得不够深,所宣称的显著调节效应很可能只是统计上的假象。

考虑一个简单回归模型:

$$y_t = \beta_0 + \beta_1 x_t + u_t. \tag{1}$$

斜率的估计量 $\mathrm{d}y/\mathrm{d}x = \widehat{\beta}_1$。若 x 和 y 的无条件方差不变,$\widehat{\beta}_1$ 决定于 x 和 y 的相关系数。所以,恒常的相关系数,或是多元回归时的偏相关系数,不允许调节效应。从这点来看,可变的相关或偏相关系数可以说是调节效应的必要条件。换言之,检验恒常的相关系数是引入交叉相乘项的先前检定。然而,拒绝了恒常相关系数的原假设,由于没有清晰的对立假设,径自采用调节效应来改善模型也免不了是一厢情愿。例如,模型参数的改变可以来自外生的冲击,无关调节变量。毕竟,非恒常相关是极广泛的概念。

二、产生虚假调节效应的模型

在线性回归中加入交叉项来挖掘数据间的调节效应易导致统计上的假象,因为变量间某些非线性关系、非恒常相关有可能被交叉项错误地捕捉。譬如,若真实模型为以下三例,交叉项很容易虚假地显著。

(一)非线性解释变量

Lubinski and Humphreys (1990) 认为,若真实模型中存在解释变量 x 的平方项,当 x 与变量 z 有线性相关时,获得显著的交叉项系数的概率上升,导致了虚假的调节效应。问题发生的原因相当直观:因为模型混淆了交叉项 xz 和 x 的平方项。当解释变量 x 和调节变量 z 的相关系数较高时,可以预见交叉项的系数几乎总是统计上显著,调节效应只是假象。这假象归咎于模型遗漏了平方项。文献里的解决方案大致如下。假设解释变量为 x 和 w 。如果被调节变量为 x ,调节变量为 z ,需要估计 y 对 $xw\,x^2xz$ 回归(若 w 为被调节变量,则考虑 w^2)。若 x^2 系数显著而 xz 系数不显著,则无调节效应。若 x^2 系数不显著,而 xz 系数显著,或者是两者均显著,则是否具有调节效应仍无定论。当 x^2 系数不显著,而 xz 系数显著时,若 z 是理论支持的唯一调节变量,据此来认定调节效应或可说是有理有据。若基于数据挖掘,z 是否为唯一的调节变量?x 是否为唯一的被调节变量?显然需要再思考。为此之故,最近的文献,例如 Daryanto (2019) 提出了以信息法则(AIC,BIC)来考虑各

种备择模型。可能的备择模型包括：

$$OLS\ y\mathrm{on}x\ z$$
$$OLS\ y\mathrm{on}x\ z\ xz$$
$$OLS\ y\mathrm{on}x\ z\ xz\ xz^2$$
$$OLS\ y\mathrm{on}x\ z\ xz\ x^2z$$
$$OLS\ y\mathrm{on}x\ x^2$$
$$OLS\ y\mathrm{on}z\ z^2$$
$$OLS\ y\mathrm{on}x\ z\ xz\ x^2\ z^2$$

这样的建议就是把调节效应的探讨带入了数据挖掘的方向。

（二）尾部相关（Tail Dependence）

尾部相关是一种特殊的非线性相关。设 X 和 Y 为两个随机变量，其分布函数分别是 F 和 G，上尾部相关（upper tail dependence）的定义是

$$\lim_{u\to 1}P\{Y > G^{-1}(u) \mid X > F^{-1}(u))=\lambda_u. \tag{2}$$

即当 X 取极端大值时，Y 也取极端大值的概率被定义为尾部相关。λ_u 越接近 1，尾部相关越大。如果随机变量 X 和 Y 间为线性相关，则尾部相关系数是零。尾部相关的情况常见于经济金融数据中，例如两个金融资产的收益率在市场正常时仅有微弱的相关性，然而在金融危机时却形成高度相关。尾部相关性的建模最直接的方式就是考虑连结函数。

假设（y_t，x_t，z_t）的真实分布具有下列特征：

(1)（y_t，x_t）的分布具有非线性相关，例如本文考虑的尾部相关性。

(2)（x_t，z_t）的分布存在线性相关。

(3)（$y_t \mid x_t$，$z_t \mid x_t$）是独立的。也就是说，给定 x_t，y_t 与 z_t 独立。

从此分布中抽样，并估计加入交叉项的回归

$$y_t=\beta_0+\beta_1 x_t+\beta_2 x_t z_t+u_t, \tag{3}$$

就容易得到统计显著的交叉项系数。这是个虚假的调节效应，因为在此真实分布中，变量 z 并不会影响变量 y 与变量 x 之间的关系。变量 z 只是与变量 x 线性相关的一个变量。事实是：任何与 x 有相关性的其他变量都有可能制造出调节效应存在的统计证据。蒙地卡罗以 Gumbel 连结函数生成具有尾部相关性数据的结果如图 1。

不难看出，z 与 x 的相关性越高，虚假的交叉相乘项的概率越大。同时，若以 Frank 连结函数来模拟，出现虚假交叉项的概率大幅降低，如图 2。原因是 Frank 连结函数没有尾部相关性。然而 Frank 连结函数的非线性相关性依然能产生虚假的调节效应，乃是因为许多连结函数能自动衍生出方差异质性。

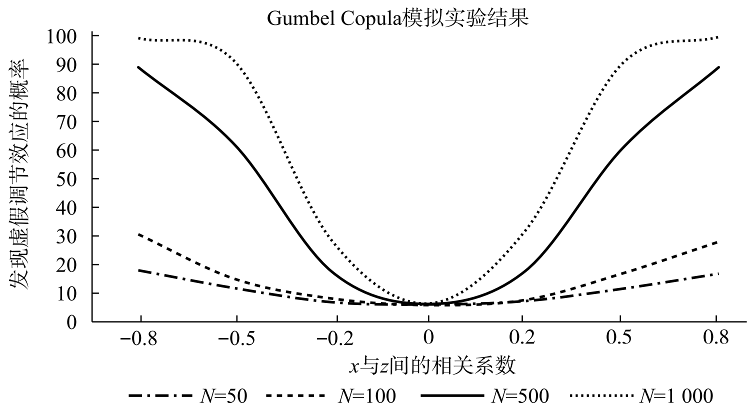

图 1 Gumbel 连结函数的模拟实验结果

注：纵坐标为交叉项显著的概率（第一类错误概率）。横坐标是 x 与 z 的线性相关系数。不同的曲线表示不同的样本量。以参数为 2 的 Gumbel 连结函数生成（y，x）的样本，且其边缘分布为标准正态分布。估计模型（3），并使用 OLS-标准误。异方差稳健标准误也会产生同样的结果。

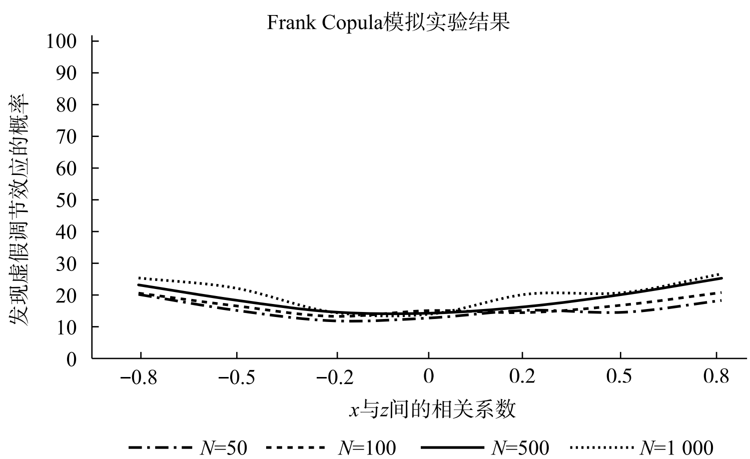

图 2 Frank 连结函数的模拟实验结果

注：纵坐标为交叉项显著的概率（第一类错误概率）。横坐标是 x 与 z 的线性相关系数。不同的曲线表示不同的样本量。以参数为 15 的 Frank 连结函数生成（y，x）的样本，且其边缘分布为标准正态分布。估计模型（3），并使用 OLS-标准误。若使用异方差稳健标准误，当样本量够大时（譬如 500 以上），虚假交叉项的概率降低到显著性水平 5%。

实证研究上也经常使用虚拟变量和 x 的交叉相乘项，常被称为异质性分析（heterogeneity analysis）。记虚拟变量为 D_i。若 D_i 取 1 的概率被 z 决定，即 $P(D_i=1)=\Phi(z_i)$，其中 Φ 为标准正态分布的分布函数，且 z 与 x 相关，模拟实验结果如表 1。事实是：若调节变量是非外生的虚拟变量，发现交叉项系数显著的概率不容忽视。

表 1　调节变量为虚拟变量的模拟实验结果

	y 与 x 线性相关	y 与 x 尾部相关
$\rho=0.2$	50	107
$\rho=0.5$	53	292
$\rho=0.8$	54	580

注：表中数字是 1 000 次实验中交叉项系数显著的次数。ρ 是 x 与 z 的线性相关系数。以参数为 2 的 Gumbel 连结函数代表尾部相关，以参数为 0.7 的 Gaussian 连结函数代表线性相关，y 与 x 的边缘分布为标准正态分布。估计模型（3），并使用 OLS-标准误。异方差稳健标准误也会产生同样的结果。

（三）方差异质性

假设

$$y_i=\alpha+\beta_i x_i+\gamma w_i+u_i, \tag{4}$$

$$\beta_i=\theta_0+v_i. \tag{5}$$

这个模型显然衍生出异方差。如果直接引入 xz 的交叉项，并使用标准的 OLS t 统计量进行交叉项的显著性检定，就容易掉入调节效应的陷阱。如表 2 显示，如果以异方差稳健标准误来检定，掉入陷阱的概率就近似于第一类误差的概率。异方差导致的虚假调节效应通过应用异方差稳健标准误即可避免，而本文后续将阐述，方程（5）随机系数的视角对调节效应的探讨有清晰的指导作用。

表 2　方差异质性与虚假调节效应

	OLS 标准误	HSK 稳健标准误
$\rho=0$	178	61
$\rho=0.2$	172	47
$\rho=0.5$	201	54
$\rho=0.8$	237	49

注：生成数据使用的参数如下：$\alpha=3$，$\gamma=0$，$\theta_0=0.5$，x_i，u_i，v_i 均服从标准正态分布。估计模型（3），表中数字是 1 000 次实验中交叉项系数显著的次数。ρ 是 x 与 z 的线性相关系数。

三、避免陷阱的辅助性检定

下列辅助性检定可用于认定合理的被调节变量，并通过排除上述产生虚假调节效应的模型来避免调节效应的陷阱。

（一）相关性恒定检验

假设基准回归为时间序列回归

$$y_t=\alpha+\beta x_t+u_t. \tag{6}$$

Wied *et al*.（2012）提出的检验可用于推断两个平稳时间序列的相关性是否恒定，进而可依此判断 y 与 x 间的调节效应是否可能存在。考虑时序样本$\{(y_t, x_t), t=1, 2, \cdots, T\}$。令$\hat{\rho}_\tau$ 为前 τ 个观测值$\{(y_1, x_1), (y_2, x_2), \cdots, (y_\tau, x_\tau)\}$的样本相关系数。于是可以计算样本相关系数集合$\{\hat{\rho}_\tau, \hat{\rho}_{\tau+1}, \cdots, \hat{\rho}_T\}$。为了检验$y_t$与$x_t$相关系数$\rho_t$在样本区间内恒定，设$H_0$：$\rho_t=\rho_0$，$t=1, \cdots, T$，$H_1$：$\exists t$，s. t. $\rho_t \neq \rho_{t+1}$。构建检验统计量

$$Q_T(x, y)=\hat{D}\max_{\tau\leqslant k\leqslant T-1}\frac{k}{\sqrt{T}}|\hat{\rho}_k-\hat{\rho}_T|, \tag{7}$$

其中 $\hat{D}$ 为样本矩的复杂的函数。给定一系列常规假设，检验统计量 $Q_T(x, y)$的渐进分布的 90%分位数为 1.223，95%分位数为 1.358，99%分位数为 1.627。

如果基准回归中包含其他控制变量 w_t

$$y_t=\alpha+\beta x_t+\gamma w_t+u_t. \tag{8}$$

此时β与y_t、x_t间的偏相关系数（partial correlation）有关。记 e_{yt} 为 y_t 对常数项和 w_t 回归所得的残差，记 e_{xt} 为 x_t 对常数项和 w_t 回归所得的残差。y_t 和 x_t 间的偏相关系数 $\rho_{xy,w}$ 等于 e_{yt}、e_{xt} 间的相关系数，因此可以利用 WKD 检验推断 e_{xt}、e_{yt} 间的相关系数是否恒定。无法拒绝 H_0 意味着调节效应存在是小概率事件。若拒绝了 H_0，可以考虑对调节效应建模，并以 x 为被调节变量。应用相同的做法来探讨是否 w 为可能的被调节变量。

若资料为横截面数据，可先将数据按被调节变量排序，然后进行 WKD 检定。

（二）随机系数生成交叉项与 LM 检定

考虑随机系数模型

$$y_i=\alpha+\beta_i x_i+\gamma w_i+u_i, \tag{9}$$

$$\beta_i=\theta_0+\theta_1 z_i+v_i. \tag{10}$$

合并方程得到

$$y_i=\alpha+(\theta_0+\theta_1 z_i)x_i+\gamma w_i+v_i x_i+u_i, \tag{11}$$

其中$(u_i+v_i x_i)$为回归模型误差项。并且，假设 v_i 与 u_i 无关且方差恒定，那么

$$\mathrm{Var}(u_i+v_i x_i)=\sigma_u^2+x_i^2\sigma_v^2, \tag{12}$$

其中 $\mathrm{Var}(u_i)=\sigma_u{}^2$，$\mathrm{Var}(v_i)=\sigma_v{}^2$。以上简单推导说明，若 z 是调节变量且 x 为被调节变量，引入交叉项的回归存在由 x^2 决定的异方差。同样地，若 w 为被调节变量，引入交叉项的回归存在由 w^2 决定的异方差。因此，调节效应的存在意味着基准回归有遗漏变量以及方差异质性。类似于检验相关性的恒定，检验基准回归的异方差也是引入交叉相乘项的先前检定。

LM 异方差检定步骤为：第一步，保存基准回归的残差 e_i。第二步，进行辅助回归：残差平方 $e_i{}^2$ 对常数项和 $x_i{}^2$ 回归，保存 R^2。则 $\mathrm{LM}= nR^2 \sim \chi_1^2$。辅助回归的解释变量也可以是 w^2，或甚至同时是 w^2 和 x^2。该检验在文献上被称为 Breusch-Pagan test。这些都是认定被调节变量的辅助检验：如果检验结果表明不存在异方差，则无法支持辅助回归的解释变量为被调节变量。如果异方差存在，则必须以异方差稳健 t 统计量来检定交叉项系数的显著性。以上述随机系数的视角对调节效应建模，相比于直接在线性回归中引入交叉项更为合理。后者等价于随机系数方程（10）省略了误差项，这意味着 β_i 与调节变量 z_i 间的关系是一个严格的线性等式，是一个过于理想的假设。

此外，建模时仅需要在随机系数方程里替换 z 成其他的调节变量便可以考虑其他调节变量可能性。该模型也不排除一个被调节变量接受两个以上调节变量的调节。在调节变量的选取上，随机系数方程应考虑容纳理论支持的各个调节变量，而非仅取某一特殊关切作调节变量。否则，建模中遗漏的调节变量亦将导致参数估计的偏误以及虚假的交叉项显著。

另一个情形是解释变量互为调节变量。假设

$$y_i = \alpha + \beta_i x_i + \gamma_i w_i + u_i, \tag{13}$$

$$\beta_i = \theta_0 + \theta_1 w_i + v_i, \tag{14}$$

$$\gamma_i = \delta_0 + \delta_1 x_i + \varepsilon_i, \tag{15}$$

代入后可以得到 xw 的交叉相乘项，而非 zx 或 zw 的交叉相乘项，即

$$y_i = \alpha + \theta_0 x_i + \delta_0 w_i + (\theta_1 + \delta_1) x_i w_i + (u_i + v_i x_i + \varepsilon_i w_i). \tag{16}$$

显然，此随机系数模型衍生了由 w^2 和 x^2 共同决定的异方差。两个特例也经常出现，即 $\theta_1 = 0$ 或 $\delta_1 = 0$。于是，y 对 x w xw 回归中 xw 的回归系数分别可能是 $(\theta_1 + \delta_1)$、δ_1 或 θ_1 的估计。实证上大量交叉项回归为此类型，但对于以上三种可能的交叉项系数的解释，少见明确的讨论。径自在线性回归中加入解释变量的乘积作交叉项将带来模型解释上的困难。

从随机系数模型的视角，交叉项回归应与异方差共存。同时检定交叉项系数的显著性需要对统计量进行异方差调整。没有异方差的交叉项回归（等于是在随机系数方程如（14）或（15）不当省略了误差项）不能支撑调节效应。另外，尾部相关，或者是非线性解释变量也能制造出虚假的调节效应，并产生异方差与交叉项回归的共存现象（如表 3）。图 1 中的样本数据就显示了方差异质性的特征。因此，为排除尾部相关和多项式回归对认定调节效应的干扰，检定尾部相关和多项式回归成了必需。

表 3　异方差与交叉项回归的共存

	尾部相关	多项式回归
$\rho = 0$	255	1 000
$\rho = 0.2$	195	1 000

（续表）

	尾部相关	多项式回归
$\rho=0.5$	116	1 000
$\rho=0.8$	67	1 000

注：ρ 是 x 与 z 间的相关系数。y 与 x 间的真实模型分别为参数为 2 的 Gumbel 尾部相关和带有 x^2 项的回归。估计模型（3）。辅助回归中的被解释变量为交叉项回归的残差的平方，解释变量为常数项和 x^2。表中数字为 1 000 次模拟实验中辅助回归的 $NR^2>3.84$（即被估模型（3）有异方差）的次数。

（三）检定尾部相关性

尾部相关检验（Reiss and Thomas，2007）可用来推断 y 与 x 间是否存在尾部相关。设 H_0 为尾部相关，检验步骤如下。第一步，把变量 y 和 x 转化为 $[0,1]$ 上的均匀分布：记 $u'=\hat{F}_n(x)$，$v'=\hat{F}_n(y)$，其中 $\hat{F}_n$ 是实证分布函数（empirical distribution function）。第二步，把 u' 和 v' 平移到左下象限：记 $u=u'-1$，$v=v'-1$，于是，u 和 v 服从 $[-1,0]$ 上的均匀分布。第三步，计算 Neyman-Pearson 检验统计量

$$T=\sum_{i=1}^{m}\log\frac{(u_i+v_i)}{c}. \tag{17}$$

$c\in(-1,0)$ 是一个从左接近 0 的阈值。m 是 u_i+v_i 超过阈值 c 的个体数量。阈值 c 的选择应使 m 约为观测样本量 n 的 10%到 15%。式（17）中要求 $u_i+v_i>c$。于是，该统计量服从随机变量为负的 Gamma 分布，其分布函数为

$$H_m(t)=e^t\sum_{i=0}^{m-1}\frac{(-t)^i}{i!},\ t\leqslant 0 \tag{18}$$

若尾部相关性存在，则调节效应有可能是虚假，显著的交叉项可能只是用于近似（approximate）y 与 x 间的尾部相关，而无关调节效应。此时，我们需要扎实的理论来支持调节效应，同时考虑多项式回归的备择模型：相比于交叉项回归，非线性解释变量更善于近似 y 与 x 间的尾部相关。若多项式回归优于交叉项回归，说明显著的交叉项仅是对非线性关系的近似，不涉及调节效应。若尾部相关性不存在，显著的交叉项也未必能支撑调节效应，需要研究交叉项回归的异方差形式来确定调节效应的合理性。

（四）多项式回归

类似于尾部相关的情况，若真实模型为多项式回归如式（19），x 和 z（或 w 和 z）的高度相关性能制造出虚假的调节效应。进行非嵌套的模型在交叉项回归和式（19）之间的选择检验，有助于避开调节效应的陷阱。

$$y_i=\alpha+\beta x_i+\gamma w_i+\delta_1 x_i^{\,2}+u_i. \tag{19}$$

非嵌套模型的选择可使用 MacKinnon 的 J test 或是 Clarke（2007）。

四、结论与建议

实证模型需要有理论的指引，调节效应属于建模的一个环节，当然也不例外。忽略应该存在的调节效应的后果是遗漏变量的偏误，因为交叉项必须是解释变量之一。换言之，交叉项在起始模型里就该出现。建模引入一般认知或是有理论支持的调节效应是为了保证主参数估计的一致性（consistent estimation）。至于交叉项的系数是否显著仅仅是次要的关切。

数据挖掘企图“证明”调节效应存在的实证研究应注意两个要点：第一，避免落入虚假调节效应的陷阱；第二，交叉项系数统计显著是研究调节效应的起点，而不是终点。图 3 建议了大致的实证策略（注意：不宜视为按步操作的攻略）。挖掘出来的调节效应需要进一步经过统计或非统计的深入探讨来建立调节效应的说服力。

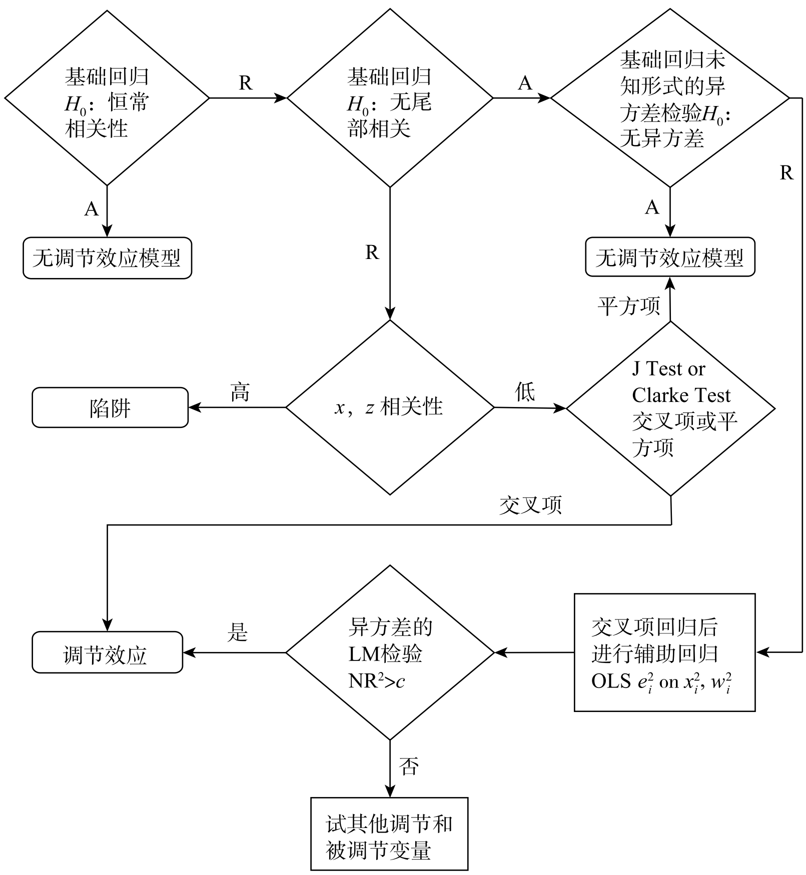

图 3　研究调节效应的实证策略

注：A 为接受 H_0，R 为拒绝 H_0。

参考文献

[1] Castro, R., G. L. Clementi, and G. MacDonald, "Investor Protection, Optimal Incentives, and Economic Growth", *The Quarterly Journal of Economics*, 2004, 119 (3), 1131-1175.

[2] Clarke, K. A., "A Simple Distribution-Free Test for Nonnested Model Selection", *Political Analysis*, 2007, 15 (3), 347-363.

[3] Daryanto, A., "Avoiding Spurious Moderation Effects: An Information-Theoretic Approach to Moderation Analysis", *Journal of Business Research*, 2019, 103, 110-118.

[4] Lubinski, D., and L. G. Humphreys, "Assessing Spurious 'Moderator Effects': Illustrated Substantivelywith the Hypothesized ('Synergistic') Relation Between Spatial and Mathematical Ability", *Psychological Bulletin*, 1990, 107 (3), 385-393.

[5] Reiss, R. D., and M. Thomas, *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, Third Edition*. Basel: Birkhäuser, 2007.

[6] Wied, D., W. Krämer, and H. Dehling, "Testing for a Change in Correlation at an Unknown Point in Time Using an Extended Functional Delta Method", *Econometric Theory*, 2012, 28 (3), 570-589.

Pitfalls of the Moderating Effect

CHIA-SHANG J. CHU* WENRUI ZHANG

(*Peking University HSBC Business School*)

Abstract Statistical significance of the interaction term in regression is often considered an evidence to support the moderating effect. There exist some pitfalls that make this statistical claim on moderating effect invalid, however. Modeling the moderating effect should be more guided by theory rather than pure data mining. If one attempts to data mine the moderating effect, some pre-tests suggested in this paper can help to avoid these pitfalls.

Keywords interaction regression, random coefficient model, moderating effect

JEL Classification C12, C21, C52

* Corresponding Author: Chia-Shang J. Chu, Peking University HSBC Business School, Shenzhen, Guangdong, 518055, China; Tel: 86-755-26032933; E-mail: cchu@phbs.pku.edu.cn.