

# 因果推断经验研究中的中介效应与调节效应

江 艇

**〔摘要〕** 中介效应分析与调节效应分析是国内经济学因果推断经验研究中被广泛采用的两种研究手段,但是存在不同程度的误用。前者的主要问题在于过度使用中介效应逐步法检验,后者主要问题是对其在因果识别中的作用阐发不足。为此,本文深入讨论了中介效应检验的偏误、如何正确开展中介效应分析,以及如何使用调节效应分析来强化因果关系论证,并针对当前的使用现状提出了相应的操作建议。

**〔关键词〕** 因果推断; 中介效应; 调节效应

**〔中图分类号〕**F064 **〔文献标识码〕**A **〔文章编号〕**1006-480X(2022)05-0120-21

## 一、引言

自20世纪八九十年代以来,经济学因果推断的经验研究范式经历了“可信性革命”(Credibility Revolution)(Angrist and Pischke, 2010),从重视函数形式的搜索和检验、重视不可观测变量方差结构的建模,转向重视研究设计,强调运用实验和准实验数据,强调寻找外生的特定冲击,强调通过挖掘研究情境的制度蕴涵和理论蕴涵而非仅依赖统计方法来进行因果关系的论证。

人类社会经济生活现象中的因果关系往往是互相缠绕、错综复杂的。当从特定原因到特定结果的因果关系被数据初步验证,研究者会进一步关心这一因果关系的作用渠道(Channel)和作用机制(Mechanism)。<sup>①</sup>分析因果关系作用渠道的出发点是,现象之间的因果关系可能包含多个逻辑环节,原因不是直接作用于结果,因此,有必要考察原因通过因果链条中的哪个或哪些中间变量影响结果,这样的分析经常被称作中介效应(Mediating Effect)分析。而分析因果关系作用机制的出发点是,一种因果关系可能不会同一不变地作用于所有个体和所有时点,因此,有必要考察因果关系的强度如何随着对象特征和现实条件的不同而产生差异,这样的分析经常被称作调节效应(Moderating Effect)分析。

新因果推断范式自然也深刻地改变了国内经济学经验研究的图景。与此同时,中介效应分析和调节效应分析这两种研究手段在近年来的经验研究论文中出现的频率越来越高,成为一种研究“时尚”,甚至是标准化操作。以《中国工业经济》为例,在最近两年发表的全部文章中,属于约简主义(Reduced-form Approach)因果推断经验研究范畴的文章一共151篇,占比63.2%;而这当中进行

**〔收稿日期〕** 2022-01-15

**〔作者简介〕** 江艇,中国人民大学经济学院副教授,经济学博士。电子邮箱:ting.jiang@gmail.com。感谢李辉文、刘瑞明、陆方文、赵勇等人极富建设性的评论和编辑部的宝贵意见,感谢何智超、王成成的助研工作,文责自负。

<sup>①</sup> 本文根据这两个词在英文文献中的惯常用法对其含义进行了严格区分,但并不是所有的文献都遵从这一区分。

了中介效应分析的论文有 85 篇,占比 56.3%;进行了调节效应分析的论文更是多达 133 篇,占比 88.1%。但不无遗憾的是,目前国内经济学研究中出现的中介效应分析和调节效应分析存在一定的盲目性和机械性,有的研究做了中介效应分析,却不问是否可信;有的研究做了调节效应分析,却不知用处何在。

鉴于此,本文尝试对这两种研究手段的工作原理和使用现状进行反思性的分析,并给出操作建议,以期引发更广泛的讨论。本文认为,目前中介效应分析的主要问题是嫁接自心理学的中介效应逐步法检验的滥用;调节效应分析的主要问题是对其在因果识别上的重要性认识不足,实证结果的理论阐发不充分。本文的结论和建议都基于一个原则:因果推断的经验研究和写作需要紧紧围绕更干净的因果识别这一首要目标。对统计手段的合理运用和对实证结果的准确阐释需要服从于这一原则。因此,本文希望达到两个目的:一是引发大家反思中介效应分析存在的科学性问题;二是引发大家重新认识调节效应分析在研究设计中的地位。最终,希望形成正确讨论因果关系作用渠道和作用机制的共识,进而提高国内经济学经验研究的可信性。

本文余下部分的结构安排如下:第二、三部分是预备知识,其中,第二部分简要介绍了何为因果推断以及因果识别的两种基本策略,第三部分介绍了中介效应分析的逐步法检验与调节效应分析的交互项模型;第四至六部分是本文的主体,其中,第四部分主要讨论中介效应逐步法检验的偏误问题,第五部分着重讨论如何开展科学的中介效应分析,第六部分转向调节效应分析,讨论其与异质性分析以及因果识别的关系;第七部分概括了这两种方法在国内经济学研究中的应用现状,并给出相应的操作建议;第八部分作简短总结。

## 二、因果识别的基本策略

定量社会科学因果推断的经验研究致力于通过大样本数据识别、估计、检验和评价社会经济生活现象之间的因果关系。用  $Y$  表示研究所关注的结果 (Outcome)、反应 (Response) 或被解释变量,用  $D$  表示有待考察的、导致结果发生的原因 (Cause)、处理 (Treatment)<sup>①</sup> 或核心解释变量。研究所关注的因果关系可以用图 1 中的基本因果模型 I 来刻画。

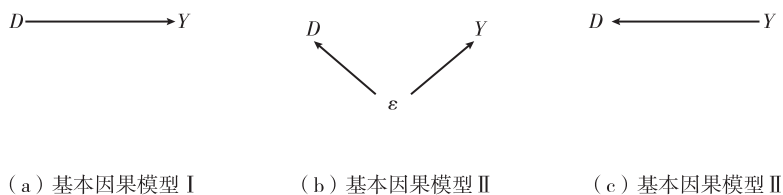


图 1 三种基本因果模型

如果从  $D$  到  $Y$  的因果关系真的存在,那么  $D$  和  $Y$  之间的相关性必然存在,反之则不然。 $D$  和  $Y$  相关这一事实可能被多个基本因果模型所合理化 (Rationalize):除了基本因果模型 I 之外,图 1 中的基本因果模型 II 表明,这种相关性可能是因为存在第三方混淆因素 (Confounding Factor) 同时影响  $D$  和  $Y$ ;

<sup>①</sup> 此处“处理”一词应作广义理解,既可能是指医学统计意义上的“治疗”或“施药”,在社会科学的研究语境中更可能是指一次实验操作、一个项目实施、一项政策干预、一次自然冲击、一个历史事件或者经济主体的一次选择行为。

基本因果模型 III 则表明,这种相关性可能是因为  $Y$  对  $D$  的反向影响 (Reverse Causality)。如果在特定的研究情境下,变量之间满足一定的假设条件,使得一个特定的因果模型没有与之竞争的、观测上等价的 (Observationally Equivalent) 的因果模型,则称这个特定的因果模型被识别,这样的假设被称作识别假设。

任何因果推断问题都包含两个部分:因果识别 (Causal Identification) 和统计推断 (Statistical Inference)。对于因果识别,所要问的是:如果拥有整个总体,是否能够确定总体因果关系? 这是社会科学理论的任务。因果识别的基本逻辑是:如果相关性不存在,则因果性不存在<sup>①</sup>;如果相关性存在且只有一种因果模型可以合理化这种相关性,则这种特定的因果性存在。对于统计推断,所要问的则是:如何从样本数据中获取关于总体因果关系的信息? 这是统计学的任务。统计推断致力于发现  $D$  和  $Y$  在样本中的相关性,并由此评估其总体相关性。

因果识别有两种基本策略。第一种基本识别策略是寻找特定的研究情境。不同的因果识别方法依赖于不同的识别假设,而不同的研究情境适用不同的识别假设。但有时候很难令人信服地论证识别假设的成立。此时研究者会尝试第二种基本识别策略——挖掘因果模型更丰富的、可验证的相关性含义 (Testable Implications),即提出这样的问题:如果从  $D$  到  $Y$  的因果关系真的存在,那么还将观测到何种现象? 不同的因果模型对新的相关性可能给出不同的预测,由此通过打破其在观测上的等价性来达到用数据验证模型的目的。

上述两种基本识别策略可总结成图 2 所示流程。一个完整的因果推断流程包含三项输入:因果问题、识别假设和数据。不同的研究情境对应不同的识别假设,从而对应不同的因果模型和因果识别方法。如果特定情境和假设下的模型和方法可以回答因果问题,则提出待估计参数,结合数据进行估计与推断,从而回答问题;如果因果问题不能被回答,则考虑更换研究情境。这就是第一种基本识别策略。由因果模型可以推导出与因果问题相关的新的可验证含义,如果这一含义被验证,则间接回答了因果问题;如果未被验证,则考虑更换研究情境。这就是第二种基本识别策略。

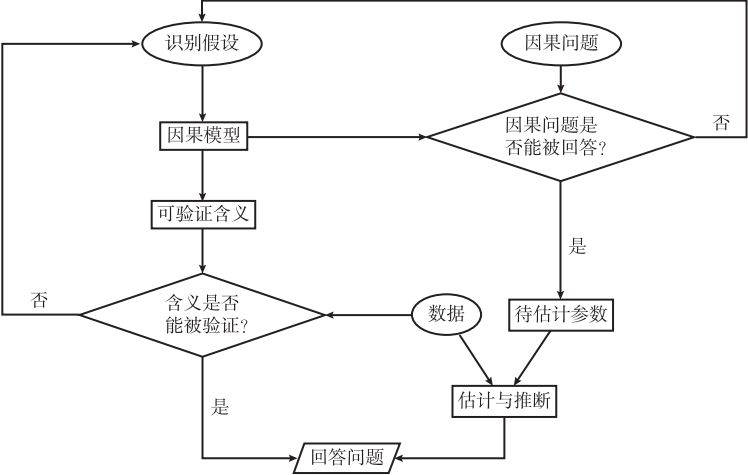


图 2 因果推断流程

三、中介效应分析与调节效应分析的现行做法

1. 中介效应的定义

中介效应是指原因通过一个或几个中间变量影响结果,这种中间变量被称作中介变量

<sup>①</sup> 有研究者认为,相关性的缺失并不能证明因果性不存在,相关性既不是因果性的必要条件,也不是其充分条件 (Bollen, 1989)。笔者则认为,当存在作用互相抵消的因果性时,不宜视作相关性不存在,而更宜理解为存在复杂的相关性。

(Mediating Factor, 或简称 Mediator)。

考察因果关系作用渠道的意义至少体现在三个方面:①能够有助于更好地认识正在考察的某项处理本身。社会科学因果推断研究经常具备的一项特征是,原因和结果的关系不那么直接而迫近,相反,它们之间的逻辑链条往往比较长,要么原因是结果的一个隐秘的原因,要么结果是结果的一个意料之外的结果(Unintended Consequence)。通过考察因果效应何以存在、如何存在,使我们更有理由相信所观测到的因果效应不是统计上的偶然,不是“第 I 类错误”,从经验上确认这项处理的内在工作方式,进而更好地理解结果的发生学。②一旦确立了这项处理经由特定渠道影响结果的因果模型,就能够更方便地判断研究结论是否可以被推广到其他总体、其他制度环境以及其他类似处理措施,换言之,可以通过考察特定渠道是否起作用来推测这种处理效应在其他情境中是否也会显现。③如果我们所关心的这项处理是来自人为设计和干预的,那么关于其作用渠道的知识和经验将有助于针对性地改进设计和干预,使其特别有利于通过特定渠道产生影响,从而提升其工作效果。

一组因果关系及其作用渠道可以用如下结构模型来刻画:

$$Y = \alpha_0 + \alpha_1 D + \varepsilon_{Y_i} \quad (1)$$

$$Y = \beta_0 + \beta_1 D + \beta_2 M + \varepsilon_{Y_i} \quad (2)$$

$$M = \gamma_0 + \gamma_1 D + \varepsilon_M \quad (3)$$

其中, $Y$ 是结果变量, $D$ 是处理变量, $M$ 是中介变量。(1)式表示 $D$ 对 $Y$ 有因果影响;(3)式表示 $D$ 对 $M$ 有因果影响;(2)式一方面表示 $M$ 对 $Y$ 有因果影响,从而建立起了 $D \rightarrow M \rightarrow Y$ 的因果链条,另一方面表示在 $M$ 之外, $D$ 还可能独立影响 $Y$ 。这组关系可以用图3直观展示。

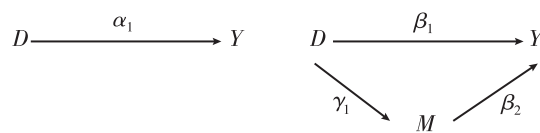


图3 中介效应示意

可称 $\alpha_1$ 为 $D$ 对 $Y$ 的总效应, $\beta_1$ 为 $D$ 对 $Y$ 的直接效应, $\beta_2\gamma_1$ 为 $D$ 对 $Y$ (经由 $M$ 中介)的间接效应,显然三者存在如下关系:

$$\alpha_1 = \beta_1 + \beta_2\gamma_1 \quad (4)$$

## 2. 中介效应检验

社会心理学家Judd和Kenny在1981年发表的论文(Judd and Kenny, 1981a)和出版的专著(Judd and Kenny, 1981b),是把中介效应分析引入社会科学项目评估的开创性工作。此后,基于线性回归的中介效应分析在心理学、流行病学、政治学、社会学、组织行为学等领域得到了广泛的应用,尤其是社会心理学研究,几乎成为必不可少的操作。本文接下来要讨论的中介效应分析不是泛指对因果关系作用渠道的探究,而是特指由Baron and Kenny(1986)提出的一整套检验流程。他们认为,中介效应的存在需要满足四个条件:① $\alpha_1$ 不为零;② $\gamma_1$ 不为零;③ $\beta_2$ 不为零;④ $\beta_1$ 为零,或至少其绝对值小于 $\alpha_1$ 。相应地,整个检验流程由四个先后步骤组成:①估计(1)式,统计上显著的 $\hat{\alpha}_1$ 意味着存在可以“被中介”的效应;②估计(3)式,统计上显著的 $\hat{\gamma}_1$ 意味着处理影响了中介;③估计(2)式,统计上显著的 $\hat{\beta}_2$ 意味着中介影响了结果;④如果 $\hat{\beta}_1$ 在统计上不显著,则意味着 $M$ 是 $D$ 与 $Y$ 关系的“完全中介”,否则意味着 $M$ 是“部分中介”。这一方法也被称为逐步法(Causal Steps Approach)。

Baron and Kenny(1986)借鉴Sobel(1982)的思想,提供了检验间接效应 $\beta_2\gamma_1$ 的方法:令 $s_{\beta_2}$ 和 $s_{\gamma_1}$ 分别表示 $\hat{\beta}_2$ 和 $\hat{\gamma}_1$ 的标准误估计, $\hat{\beta}_2\hat{\gamma}_1$ 的标准误估计可以近似为 $s_{\beta_2\gamma_1} = \sqrt{\hat{\beta}_2^2 s_{\gamma_1}^2 + \hat{\gamma}_1^2 s_{\beta_2}^2 + s_{\beta_2}^2 s_{\gamma_1}^2}$ ,由此构造统计量 $\hat{\beta}_2\hat{\gamma}_1/s_{\beta_2\gamma_1}$ ,渐进服从标准正态分布。另外一些文献,如Shrout and Bolger(2002),建议使

用自抽样(Bootstrap)方法得到间接效应的自抽样分布,进而计算其自抽样标准误和自抽样置信区间。

实际上,因为  $\alpha_1 - \beta_1 = \beta_2 \gamma_1$ , 检验间接效应等价于检验  $\alpha_1$  与  $\beta_1$  是否相等。注意这是跨方程的系数差异检验,但可以通过堆叠数据的方式方便地实现。

在图4中,左侧的表格是原始数据示例,将原始数据复制一份得到右侧的表格,然后构造虚拟变量  $W$  来区分每份数据,再构造三个新的交互项  $D(1-W)$ 、 $DW$  以及  $MW$ 。

id	Y	D	M	id	Y	W	D (1-W)	DW	MW
1	$y_1$	$d_1$	$m_1$	1	$y_1$	0	$d_1$	0	0
2	$y_2$	$d_2$	$m_2$	2	$y_2$	0	$d_2$	0	0
3	$y_3$	$d_3$	$m_3$	3	$y_3$	0	$d_3$	0	0
1	$y_1$	$d_1$	$m_1$	1	$y_1$	1	0	$d_1$	$m_1$
2	$y_2$	$d_2$	$m_2$	2	$y_2$	1	0	$d_2$	$m_2$
3	$y_3$	$d_3$	$m_3$	3	$y_3$	1	0	$d_3$	$m_3$

图4 堆叠数据示例

将  $Y$  回归在  $W$ 、 $D(1-W)$ 、 $DW$  以及  $MW$  上:

$$E(Y|\cdot) = \delta_0 + \delta_1 W + \delta_2 D(1-W) + \delta_3 DW + \delta_4 MW = \begin{cases} \delta_0 + \delta_2 D, W=0 \\ (\delta_0 + \delta_1) + \delta_3 D + \delta_4 M, W=1 \end{cases} \quad (5)$$

比较(1)式、(2)式与(5)式可见,  $\delta_2 = \alpha_1$ ,  $\delta_3 = \beta_1$ , 因此,在(5)式的回归中检验  $\delta_2$  与  $\delta_3$  是否相等,等价于检验  $\alpha_1$  与  $\beta_1$  是否相等。<sup>①</sup>

### 3. 调节效应的定义与常见情形

调节效应是指原因对结果的影响强度会因个体特征或环境条件而异,这种特征或条件被称作调节变量(Moderating Factor,或简称 Moderator),如图5所示。交互项模型是对调节效应进行建模的主要方式。

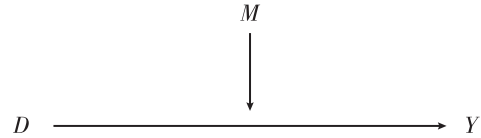


图5 调节效应示意

(1)两个处理变量的情形。先看一种较少见的情形。构成交互项的两个解释变量都是核心解释变量:

$$Y = \beta_0 + \beta_1 D_1 + \beta_2 D_2 + \beta_3 D_1 \times D_2 + \varepsilon \quad (6)$$

两个核心解释变量对  $Y$  的边际效应分别为:

$$\frac{\partial E(Y|\cdot)}{\partial D_1} = \beta_1 + \beta_3 D_2, \quad \frac{\partial E(Y|\cdot)}{\partial D_2} = \beta_2 + \beta_3 D_1 \quad (7)$$

此时,较少说两个核心解释变量互为调节,而是说该模型适用于考察两个变量对  $Y$  影响的互补性( $\beta_3 > 0$ )或替代性( $\beta_3 < 0$ )。

(2)一个处理变量的情形。在大多数时候,一项因果推断研究只关注一个处理变量,仍记作  $D$ ,将调节变量记作  $M$ ,则有:

$$Y = \beta_0 + \beta_1 D + \beta_2 M + \beta_3 M \times D + \varepsilon \quad (8)$$

$$\frac{\partial E(Y|\cdot)}{\partial D} = \beta_1 + \beta_3 M \quad (9)$$

此时称  $D$  对  $Y$  的因果影响受到  $M$  的调节:如果  $\beta_3 > 0$ ,则  $D$  对  $Y$  的正面影响随着  $M$  的增大而

<sup>①</sup> 唯一需要注意的是系数估计标准误要聚类到 id 层面。

增强(或负面影响随着  $M$  的增大而减弱);如果  $\beta_3 < 0$ , 则  $D$  对  $Y$  的正面影响随着  $M$  的增大而减弱(或负面影响随着  $M$  的增大而增强)。

如果  $\hat{\beta}_3$  在统计上显著, 则称观测到了显著的调节效应。但  $\hat{\beta}_1$  的大小和统计显著性是否具有经济意义, 要视具体的研究情境而定, 因为此时  $\beta_1$  的含义是当  $M=0$  时  $D$  对  $Y$  的因果效应。很有可能当  $M=0$  时  $D$  对  $Y$  影响不显著, 才是符合理论预期的。

除了报告系数的估计结果之外, 通常还会根据需求, 报告因果效应的一个或几个估计值(以及标准误、置信区间、显著性检验的  $p$  值等)。例如,  $\hat{\beta}_1 + \hat{\beta}_3 \bar{M}$  (对于样本中的平均观测个体,  $D$  对  $Y$  的因果效应有多大), 或者  $\hat{\beta}_1 + \hat{\beta}_3 M_{max}$ 、 $\hat{\beta}_1 + \hat{\beta}_3 M_{min}$  等。也可以用图形的方式直观展示, 如图 6 所示。在这个示意图中,  $\hat{\beta}_1 < 0$ ,  $\hat{\beta}_3 > 0$ , 实线表示因果效应的估计值, 虚线表示给定置信水平的置信区间。当  $M < m_2$  时,  $D$  对  $Y$  的因果效应是负向的; 当  $M < m_1$  时, 这种负向效应在统计上显著; 当  $M > m_2$  时,  $D$  对  $Y$  的因果效应是正向的; 当  $M > m_3$  时, 这种正向效应在统计上显著。

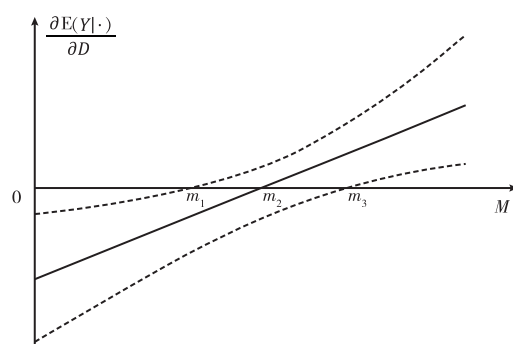


图 6 展示调节效应的估计结果

有时会对调节变量进行中心化(去均值)处理, 将(8)式改造成如下等价形式:

$$Y = \delta_0 + \delta_1 D + \delta_2 (M - \bar{M}) + \delta_3 (M - \bar{M}) \times D + \varepsilon \quad (10)$$

进行这种处理的理由并不是有些文献所声称的为了避免变量之间的多重共线性, 而仅仅是为了让一次项系数  $\delta_1$  具有更直观的含义:

$$\frac{\partial E(Y|\cdot)}{\partial D} = \delta_1 + \delta_3 (M - \bar{M}) \quad (11)$$

可知, 此时  $\delta_1$  为平均观测个体 ( $M = \bar{M}$ ) 的  $D$  对  $Y$  的因果效应。

(3) 调节变量为虚拟变量的情形。当调节变量  $M$  为虚拟变量时, 还可以把交互项模型等价地理解为分组回归:

$$E(Y|\cdot) = \begin{cases} \beta_0 + \beta_1 D, & M = 0 \\ (\beta_0 + \beta_2) + (\beta_1 + \beta_3) D, & M = 1 \end{cases} \quad (12)$$

换言之, 分组回归下  $D$  对  $Y$  的因果效应的组间异质性可以通过交互项模型来检验, 即表现为交互项  $M \times D$  系数估计的统计显著性。这种检验总是有必要的, 因为组间异质性不能诉诸于直观判断。例如,  $D$  对  $Y$  的因果效应在  $M=0$  组是不显著的, 在  $M=1$  组是显著的, 但它们之间的差异很可能是显著的。

有时为方便展示结果, 对(8)式进行如下等价变换:

$$Y = \gamma_0 + \gamma_1 M + \gamma_2 M \times D + \gamma_3 (1 - M) \times D + \varepsilon \quad (13)$$

此时  $\gamma_2$  和  $\gamma_3$  分别是  $M=1$  组和  $M=0$  组的  $D$  对  $Y$  的因果效应。

因此, 当调节变量为虚拟变量时, 有三种呈现结果的方式。第一种是报告(8)式的结果, 这种方式的好处是直接显示了  $D$  对  $Y$  的因果效应的组间异质性, 缺陷是  $M=1$  组的因果效应没有直接显示。第二种是分组报告(12)式的结果, 这种方式的好处是直接显示了这两个组的因果效应, 缺陷是



因果效应的组间异质性检验还需要额外通过(8)式来实现。<sup>①</sup>第三种是报告(13)式的结果,这种方式的好处也是直接显示了这两个组的因果效应,因果效应的组间差异虽然没有直接显示,但是可以方便地通过检验  $H_0: \gamma_2 = \gamma_3$  来实现。

(4) 处理变量和调节变量均为虚拟变量的情形。当处理变量  $D$  和调节变量  $M$  均为虚拟变量时,交互项系数有一种方便的理解:处理组( $D=1$ )与控制组( $D=0$ )的结果均值的组间差异之差异,简称双重差分(Difference in Differences)。对于(8)式,易知:

$$E(Y|\cdot) = \begin{cases} \beta_0 + \beta_1 + \beta_2 + \beta_3, D=1, M=1 \\ \beta_0 + \beta_2, D=0, M=1 \\ \beta_0 + \beta_1, D=1, M=0 \\ \beta_0, D=0, M=0 \end{cases} \quad (14)$$

$$\beta_3 = [E(Y|D=1, M=1) - E(Y|D=0, M=1)] - [E(Y|D=1, M=0) - E(Y|D=0, M=0)] \quad (15)$$

例如, $Y$ 表示工资水平, $D$ 表示是否上大学, $M$ 表示是否为男性,则 $\beta_3$ 表示男性的上大学回报率与女性的上大学回报率的差异。

特别地,当 $D$ 为处理实施后的虚拟变量, $M$ 为是否最终接受处理的虚拟变量,这种特殊的调节效应模型就是双重差分模型。反过来说,无论 $D$ 和 $M$ 是离散变量还是连续变量,无论其变动性是截面维度还是时间维度,交互项系数都应该在双重差分的意义下去理解。

这里涉及的一个具体问题是,如何在双重差分的意义下表述交互项系数的经济含义。当 $D$ 和 $M$ 均为虚拟变量时,含义的表述比较方便: $\beta_3$ 表示处理的实施对结果的影响。当 $M$ 为连续变量时,意味着所有个体都接受了处理,只是处理的强度(Treatment Intensity)有所不同,此时含义可以表述为:当 $M$ 增加一个标准差 $\sigma_M$ 时,处理的实施对结果的影响会提高 $\sigma_M \cdot \beta_3$ 。当 $D$ 和 $M$ 均为连续变量时,表述尤其需要小心。Rajan and Zingales(1998)使用跨国行业层面数据研究金融市场的发展是否影响产业增长。 $D$ 为反映一国金融发展水平的指标, $M$ 为反映某一行业外部融资依存度的指标, $Y$ 为该国该行业增加值的增长率。 $\beta_3 > 0$ ,表明更加依赖于外部融资的行业,在金融市场发展水平越高的国家增长得更快。研究结果的经济含义是:在样本中,典型的高外部融资依存行业是 $M$ 居于75分位的机床业,典型的低外部融资依存行业是 $M$ 居于25分位的饮料业,典型的高金融发展水平国家是 $D$ 居于75分位的意大利,典型的低金融发展水平国家是 $D$ 居于25分位的菲律宾,因此,相较于菲律宾,在意大利机床业的增长比饮料业要快 $\hat{\beta}_3 \cdot (M_{75} - M_{25}) \cdot (D_{75} - D_{25})$ ,然后将这一数字与 $Y$ 的样本均值进行比较,表明其足够大以至不可忽略。

## 四、对中介效应检验的反思

### 1. 心理学对中介效应检验的反思

Judd and Kenny(1981a, 1981b)在引入中介效应分析之初曾明确指出,必须认识到中介效应分析的本质是一种相关性分析。具体而言,用线性回归方法考察(1)式—(3)式会存在两个问题:统计检验的功效较低,效应的估计可能存在偏误(Judd and Kenny, 1981a)。低统计功效来源于两个

<sup>①</sup> 一个技术细节是,如果(12)式中包含控制变量,那么对应的(8)式中必须包含所有控制变量与调节变量的交互项,才能保证(8)式与(12)式估计结果的等价性。

方面:一是处理变量和中介变量是高度相关的,因此,在以结果变量为被解释变量的回归,即(2)式中,两者可能高度共线性,从而增大系数估计的标准误,降低统计显著性,也就是说,如果无法拒绝 $\beta_1 = 0$ 的原假说(即所谓“完全中介”),那么可能是因为统计功效太低;二是中介变量和结果变量都可能存在测量误差,以其为被解释变量时,测量误差会导致统计功效降低。与此相比,估计偏误问题更为关键:①虽然有可能通过随机实验的方式保证处理的外生性,但几乎总是无法干预中介变量,因此,可能存在既影响中介变量又影响结果变量的混淆因素,如果这种因素在分析中被遗漏,就有可能导致(2)式中的系数估计产生偏误,这一问题即便在随机实验研究中也无法避免。②中介变量的测量误差倾向于使(2)式中 $\beta_2$ 的估计产生趋零偏误,进而使得 $\beta_1$ 被高估。③中介变量与结果变量可能是互为因果的,这也会导致(2)式产生估计偏误,偏误的方向取决于双向因果的符号。

Kenny et al.(1998)指出,逐步法中的第1步不是必须的。如果 $\alpha_1$ 等于零,则意味着间接效应与直接效应方向相反。例如, $D$ 表示儿童的智商, $Y$ 表示儿童的情绪稳定性,通常智商越高的儿童情绪越稳定,但高智商的儿童可能更容易受到同学的欺负,而被欺负将导致其情绪不稳定。此时智商对情绪稳定的间接效应是负向的,与直接效应方向相反(Judd and Kenny, 2010)。又如, $D$ 表示负面新闻报道, $Y$ 表示民众的集体参与。一方面,负面新闻报道使得民众认识到社会后果的重要性从而鼓励了集体参与;另一方面,负面新闻报道削弱了民众对政府的信任从而降低了集体参与。这两种中介所起的作用相反,可能互相抵消(Hayes, 2009)。因此,Kenny et al.(1998)认为,确认中介效应存在的关键是逐步法中的第2步和第3步。

鉴于间接效应 $\beta_2\gamma_1$ 的统计显著性与 $\beta_2$ 或 $\gamma_1$ 的统计显著性没有必然联系,逐步法中的第2步或第3步检验未通过并不意味着间接效应不存在,因此,更多的学者认为,第2步和第3步也不是必须的,唯一重要的步骤是检验 $\beta_2\gamma_1 = \alpha_1 - \beta_1 \neq 0$ (Hayes, 2009)。但Judd and Kenny(2010)告诫到,将中介效应分析退化为单一的统计检验,可能会使研究者误以为中介效应是可以通过纯粹的统计手段来识别的,要主张中介效应的存在,只需要检验其在统计上显著即可,而不再关心中介效应模型背后的一系列因果识别假设。

此外,Hayes(2022)批评了完全中介和部分中介的提法。首先,在统计上发现了完全中介,并不意味着对 $D$ 影响 $Y$ 的作用渠道的探究就可以终止了。反过来说,有可能存在多个中介,它们在单独的检验中分别都是统计意义上的完全中介,那么完全中介就是一个空洞的术语。其次,如果在统计上发现了部分中介,则意味着至少还有一些中介尚未被当前模型所刻画,当前模型在一定程度上是误设的,因此,部分中介也不是一个有价值的结论。

## 2. 中介效应检验的偏误

从前文的介绍中可以看出,心理学中的中介效应检验与经济学中的因果推断研究的“格格不入”之处在于,后者的研究主题是考察处理变量对结果变量的因果影响,如何选择合适的研究情境,运用经济学理论论证关键的识别假设,提高研究手段的因果识别力(Identification Power),是研究者主要致力的方向;而前者的聚焦点则是处理变量经由中介变量影响结果变量这一间接效应本身。因此,在前者的研究中,关键的识别假设被有意无意地认为是不证自明的,因果识别似乎“在场”,却分明“缺席”。考虑到一方面国际学术界正在对不严谨的中介效应分析的泛滥进行反思(Bullock et al., 2010),另一方面中介效应检验在国内经济学界大有“攻城略地”之势,笔者认为,迫切的任务是直观地展示中介效应检验的核心缺陷所在。

假定 $D$ 是一种随机处理,即 $E(\varepsilon_Y D) = 0$ 和 $E(\varepsilon_M D) = 0$ 成立,因此,(1)式和(3)式的普通最小二乘估计分别能够得到 $\alpha_1$ 和 $\gamma_1$ 的一致估计:



$$\hat{\alpha}_1 \xrightarrow{p} \alpha_1, \hat{\gamma}_1 \xrightarrow{p} \gamma_1 \quad (16)$$

在(2)式的普通最小二乘估计中,根据 Frisch-Waugh-Lovell 定理,有:

$$\hat{\beta}_2 \xrightarrow{p} \frac{\text{Cov}(Y, \varepsilon_M)}{\text{Var}(\varepsilon_M)} = \frac{\text{Cov}(\beta_0 + \beta_1 D + \beta_2 M + \varepsilon_{Y_2}, \varepsilon_M)}{\text{Var}(\varepsilon_M)} = \beta_2 + \frac{\text{Cov}(\varepsilon_{Y_2}, \varepsilon_M)}{\text{Var}(\varepsilon_M)} \quad (17)$$

其中,第二个等号用到  $\text{Cov}(\varepsilon_M, D) = 0$  以及  $\text{Cov}(M, \varepsilon_M) = \text{Var}(\varepsilon_M)$ 。

将(3)式代入(2)式,有:

$$Y = (\beta_0 + \beta_2 \gamma_0) + (\beta_1 + \beta_2 \gamma_1) D + (\varepsilon_{Y_2} + \beta_2 \varepsilon_M) \quad (18)$$

可知  $\alpha_0 = \beta_0 + \beta_2 \gamma_0, \alpha_1 = \beta_1 + \beta_2 \gamma_1, \varepsilon_{Y_1} = \varepsilon_{Y_2} + \beta_2 \varepsilon_M$ 。特别地,有:

$$\hat{\beta}_1 + \hat{\beta}_2 \hat{\gamma}_1 \xrightarrow{p} \beta_1 + \beta_2 \gamma_1 \quad (19)$$

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 + \gamma_1 (\beta_2 - \hat{\beta}_2) = \beta_1 - \gamma_1 \frac{\text{Cov}(\varepsilon_{Y_2}, \varepsilon_M)}{\text{Var}(\varepsilon_M)} \quad (20)$$

因此,只有当  $\frac{\text{Cov}(\varepsilon_{Y_2}, \varepsilon_M)}{\text{Var}(\varepsilon_M)} = 0$ ,即不存在同时影响  $M$  和  $Y$  的不可观测因素、不存在从  $Y$  到  $M$  的反向因果,且  $M$  不存在测量误差时,简言之,只有当  $M$  也可以外生随机干预时,  $\beta_1$  和  $\beta_2$  的普通最小二乘估计才是一致的。

总之,如果处理是随机的但中介不是随机的,那么通过估计(1)式可以得到处理对结果的因果效应,通过估计(3)式可以得到处理对中介的因果效应,但(2)式的最小二乘估计是不可靠的,无法得到处理对结果的直接因果效应( $\beta_1$ ),也无法得到中介对结果的因果效应( $\beta_2$ ),从而无法得到处理对结果的间接因果效应( $\beta_2 \gamma_1$ )。

对图3稍作补全,就很好理解上述结论了。当进行中介效应检验时,隐含的关于数据生成过程的假设如图7(a)所示,但真实的数据生成过程很可能如图7(b)所示。Robins and Greenland(1992)在鲁宾因果模型(Rubin's Causal Model)框架下证明,仅有随机处理,无法识别直接效应和间接效应;即使没有延滞效应(Carryover Effect)的交叉试验(Crossover Trial)是可行的<sup>①</sup>,也无法识别;只有当研究者可以对中介变量进行外生干预时,才有可能识别。<sup>②</sup>

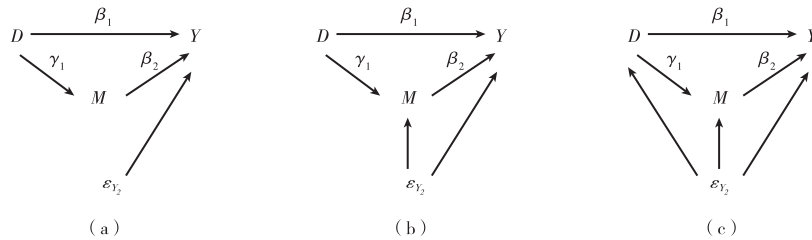


图7 三种不同的数据生成过程

为了验证理论推导的结论,考虑如下数据生成过程:

$$D \sim N(0, 1) \quad (21)$$

$$e \sim N(0, 1) \quad (22)$$

① 即指可以对同一个被试先后施加两种处理状态,从而观测其接受处理和未接受处理时的结果,并且两种处理状态的先后顺序不会影响被试的结果。

② 以二值中介变量为例,这种干预意味着研究者可以在  $M = 1$  的处理组个体 ( $D = 1, M = 1$ ) 和控制组个体 ( $D = 0, M = 1$ ) 中分别随机选取一部分,设定其  $M = 0$ 。

$$M = D + e + \varepsilon'_M, \quad \varepsilon'_M \sim N(0, 1) \quad (23)$$

$$Y = D + M + \rho e + \varepsilon'_{Y_2}, \quad \varepsilon'_{Y_2} \sim N(0, 1) \quad (24)$$

可知  $D$  是一种随机处理, 与  $\varepsilon'_M$  以及  $\varepsilon'_{Y_2}$  均独立。 $D$  对  $Y$  的直接效应  $\beta_1 = 1$ ,  $M$  对  $Y$  的效应  $\beta_2 = 1$ ,  $D$  对  $M$  的效应  $\gamma_1 = 1$ ,  $D$  对  $Y$  的间接效应  $\beta_2 \gamma_1 = 1$ ,  $D$  对  $Y$  的总效应  $\alpha_1 = 2$ 。 $e$  是研究者观测不到的同时影响  $M$  和  $Y$  的因素,  $\rho$  刻画了由  $e$  引起的  $M$  和  $Y$  之间的相关性。 $\varepsilon_{Y_2} = \rho e + \varepsilon'_{Y_2}$ ,  $\varepsilon_M = e + \varepsilon'_M$ , 因此,  $\frac{\text{Cov}(\varepsilon_{Y_2}, \varepsilon_M)}{\text{Var}(\varepsilon_M)} = \rho/2$ 。

通过以上数据生成过程, 随机生成样本量为 100 的样本, 然后用普通最小二乘法分别估计 (1) 式—(3) 式, 如此重复 1000 次, 得到相关系数估计量的经验分布。先来看  $\rho = 1$  时的结果。

由图 8(a) 可见, 处理对结果的总效应估计和处理对中介的效应估计没有出现偏误。由图 8(b) 可见, 处理对结果的直接效应估计集中在 0.5 附近, 偏离了其真实值 1; 处理对结果的间接效应估计集中在 1.5 附近, 也偏离了其真实值 1。这种偏误正是由中介对结果的效应估计出现偏误所致。也就是说, 当不可观测变量引起  $M$  和  $Y$  正相关时, 直接效应会被低估, 间接效应会被高估, 偏误的方向和大小与理论推导的结论相符。

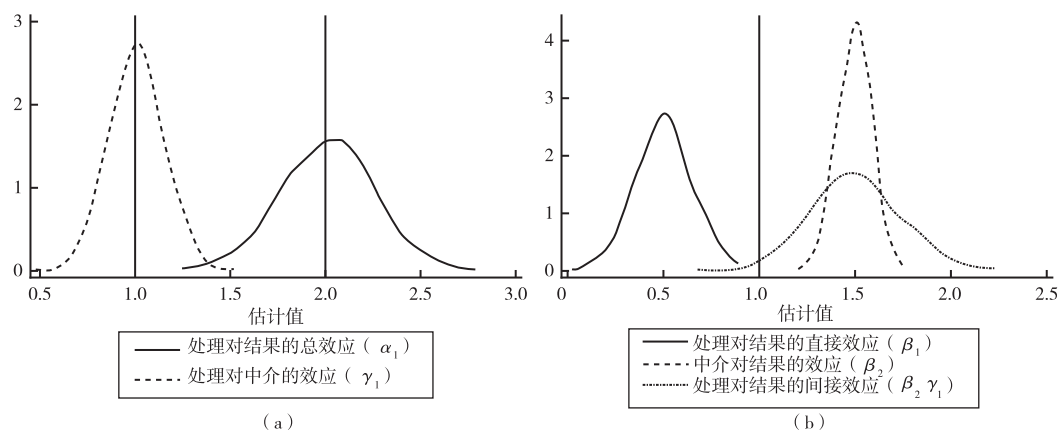


图 8 数据模拟结果: 部分中介,  $\rho = 1$

基于同样的理由, 即使不寻求对直接效应和间接效应的无偏估计, 转而采取一种更保守的因果推断立场: 通过观察  $Y$  对  $D$  的回归中加入  $M$  后  $D$  的系数大小是否发生显著改变, 也就是通过观察 (或正式地统计检验)  $\beta_1$  估计的绝对值是否显著小于  $\alpha_1$  估计的绝对值来判断是否存在中介效应, 这种研究手段也可能导致错误的结论。例如, 把 (24) 式改为:

$$Y = M + \rho e + \varepsilon'_{Y_2}, \quad \varepsilon'_{Y_2} \sim N(0, 1) \quad (25)$$

此时面临的是完全中介情形, 并令  $\rho = -2$ 。由图 9(a) 可见, 直接效应估计集中在 1 附近, 偏离了其真实值 0; 间接效应估计集中在 0 附近, 偏离了其真实值 1。事实上, 如果在每次模拟中都采用堆叠回归的方式对原假说  $H_0: \alpha_1 = \beta_1$  进行检验, 1000 次模拟中只有 69 次可以在 10% 的显著性水平上拒绝该原假说, 也就是说, 只有 6.9% 的概率可以正确地发现中介效应的存在。

相反, 如果把 (24) 式改为:

$$Y = D + \rho e + \varepsilon'_{Y_2}, \quad \varepsilon'_{Y_2} \sim N(0, 1) \quad (26)$$

此时面临的是无中介效应的情形, 并令  $\rho = 2$ 。但是由图 9(b) 可见, 直接效应估计集中在 0 附近, 偏离了其真实值 1; 间接效应估计集中在 1 附近, 偏离了其真实值 0。事实上, 如果在每次模拟

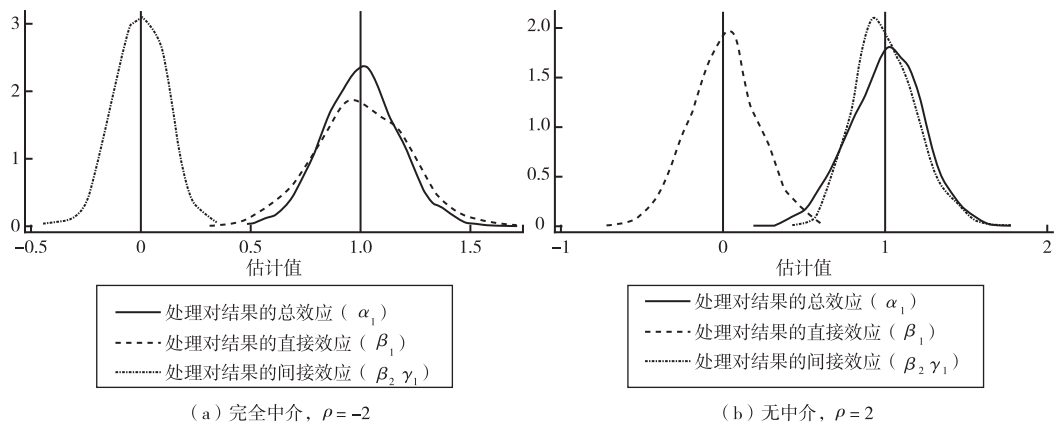


图9 数据模拟结果

中同样对原假说  $H_0: \alpha_1 = \beta_1$  进行检验,1000 次模拟全部可以在 1% 的显著性水平上拒绝该原假说,也就是说,总是得到完全中介的错误结论。

在以处理观测性数据为主的经济学因果推断研究中,研究者面临着更富有挑战性的处境:真实的数据生成过程很可能如图 7(c) 所示。此时处理变量的生成方式不再来自于研究者外生的实验干预,而来自于研究对象的主动选择,因此,研究者在研究设计上的努力主要聚焦在如何解决处理变量的内生性问题,中介变量的内生性问题不得不成为等而次之的问题。或者说,解决处理变量的内生性并探究其对结果的影响,同解决中介变量的内生性并探究其对结果的影响,应该成为两项独立研究各自的主题,以确保每项研究都聚焦在一个核心解释变量上。除非能够从理论上证明并不存在同时影响中介变量和结果变量的混淆因素,或者能够良好地定义、充分地穷举和准确地测度这些混淆因素并且方便地将其作为控制变量放入(2)式中,否则这种回归并不能产生任何增进我们理解  $D \rightarrow M \rightarrow Y$  这一因果链条的有益知识。

## 五、经济学中的中介效应分析

### 1. 中介效应检验为何少见

上文的讨论表明,中介效应检验的适用前提是,识别  $D$  对  $M$  和  $Y$  的因果关系比较容易,同时识别  $M$  对  $Y$  的因果关系也比较容易。而对于观测性数据研究,真实的数据生成过程纷繁复杂,找到合适的研究情境来研究  $D$  对  $Y$  的因果关系已属不易,研究中介效应更是困难,这就是为什么中介效应检验历来在经济学经验研究文献中很少见的主要原因。

即使不考虑中介效应检验的内生性偏误,大多数开展此类检验的研究往往发现, $D$  对  $Y$  的总效应中,除了通过  $M$  发生的间接效应之外,有相当一部分是直接效应,也就是说(2)式的估计中, $\beta_1$  往往显著不为零, $M$  是“部分中介”。这是由社会经济现象的复杂性所决定的,原本是可以预期到的,但将其作为正式研究结论的一部分却并非值得称道之事,实际上暴露出研究者对  $D$  究竟如何影响  $Y$ ,有相当一部分是不清楚的。

中介效应检验逐步法的关键一步,是把中介变量  $M$  作为控制变量加入  $Y$  对  $D$  的回归中。但在观测性数据研究中,加入控制变量的目的主要是为了解决由控制变量所造成的处理的内生性,研究者指望,在控制变量取值相同的子总体(Subpopulation)内, $D$  的变动不再受到不可观测的选择性的影

响,可以近似看作外生的随机变动。所以,好的控制变量首先应该是一个处理前变量(Pretreatment Variable),其发生在 $D$ 之前,既影响 $D$ ,又影响 $Y$ 。相反,发生在 $D$ 之后的控制变量是“坏”的控制变量,而控制中介变量就是一种典型的“坏控制”(Bad Control)(Angrist and Pischke, 2009)。即使在简单随机实验研究中,这一结论也成立。此时研究者并不需要控制变量来帮助识别, $Y$ 对 $D$ 的回归可以得到 $\alpha_1$ 的一致估计,但控制 $M$ 之后,反而得不到 $\beta_1$ 的一致估计。

中介效应检验不可靠,并不意味着不研究因果关系的作用渠道。一种常见的做法是,提出一个或几个中介变量 $M$ ,这些变量和 $Y$ 的因果关系在理论上比较直观,在逻辑和时空关系上都比较接近,以至于不必采用正式的因果推断手段来研究从 $M$ 到 $Y$ 的因果关系;然后仅看 $D$ 对 $M$ 的影响,即只考察(1)式和(3)式,而不考察(2)式,从而避免正式区分出在间接效应之外是否还有无法解释的直接效应。这样的例子比比皆是。例如,Dell(2011)通过地理回归断点设计方法,发现16—19世纪秘鲁的强制徭役制度( $D$ )导致当代居民家庭消费水平降低以及儿童发育迟缓( $Y$ );在随后的“持续影响的作用渠道”(Channels of Persistence)分析中,将被解释变量依次替换为土地所有权、教育和道路等公共品供给、劳动供给以及市场参与等消费的直接决定因素( $M$ ),即视为达到了检验渠道的目的,至于这些 $M$ 如何影响 $Y$ 、 $D$ 经由 $M$ 影响 $Y$ 的效应有多大,以及 $D$ 是否在影响 $M$ 之外还会直接影响 $Y$ ,就不再着墨了。

## 2. 如何理解经济学中的中介效应检验

确实有少量研究在一定程度上借鉴了中介效应检验逐步法的思路,考察 $Y$ 对 $D$ 的回归中加入中介变量 $M$ 后 $D$ 的系数估计绝对值是否减少,以此论证 $M$ 是否为 $D$ 作用于 $Y$ 的渠道。不过一旦认识到这种做法的固有缺陷,就不会过于强调论证力度,而只是将其视作某种试探性的证据。下面举两个例子。

第一个例子来自Alesina and Zhuravskaya(2011)。这是一项跨国研究,构造了每个国家的族群分散指数,发现分散指数越高的国家,政府的治理质量越低。用周边国家的族群分布来预测每个国家的族群分布(相同的族群倾向于靠近彼此的国界),从而构造出一国分散指数的工具变量,识别了族群分散对政府治理质量的因果影响。然后,提出了因果关系的三个作用渠道:人际信任、国家凝聚力和族群投票行为,同样用工具变量方法识别了分散指数对这三个中介变量的因果影响。在“渠道检验”(Test of the Potential Channels)中指出,要使这三个变量能够成为作用渠道,除了要和分散指数相关之外,还必须和政府质量相关,在政府质量对分散指数的回归中控制这些变量后,分散指数的系数估计绝对值应该会下降。结果表明,当控制人际信任后,分散指数的系数估计绝对值确实大幅下降,而且人际信任的系数显著为正;当继续控制另外两个中介变量后,分散指数的系数估计不再发生变化,两个中介变量的系数不显著。他们认为,这一结果意味着人际信任是将族群分散和政府质量联系起来的重要渠道,但不足以断定国家凝聚力和族群投票行为就不是额外的渠道,因为这三个变量是高度相关的,因此很难把每个渠道彼此分离。由此看出,他们在识别因果关系的作用渠道时,仍然更多地依赖于中介变量对处理变量的回归结果,而只是将诸如(2)式的回归结果作为一个旁证。

第二个例子来自Persico et al.(2004)。利用英国和美国青少年大型调查数据发现,在控制人口统计学、家庭背景和个人健康智力等特征后,青少年时期的身高与成年的工资水平显著正相关。随后提出了一系列潜在的作用渠道:职业选择、自尊、社交活动、能力测试。他们没有把这些中介变量对青少年身高进行回归,而是将其作为控制变量依次控制在工资水平对青少年身高的回归中。当控制职业选择和自尊后,青少年身高的系数估计没有发生大的改变,因而认为这两个变量的中介作用有限。当控制社交活动数量后,青少年身高的系数估计值下降了近40%且不再显著,而社交活动的系数估计显著为正。控制能力测试成绩也得到了类似的结果。但他们认为必须谨慎地解释这些结果。社交活动本身是一种主动选择,而该文中并没有对这一选择进行正式的建模,因此不能认为回归结果反

映了社交活动对工资的因果效应。在解释能力测试与工资水平的显著正相关时,也十分小心地用辅助证据来论证这种相关性不是出于先天智商的混淆影响,而确实可能说明身高会影响青少年经历。<sup>①</sup>

无论如何,在这两个例子中,都没有估计间接效应的大小并对其统计显著性进行正式检验。这仍然涉及对中介效应分析应该服务于何种目的的理解:究竟是在度量的意义上还是在解释的意义上考察中介效应?如果是前者,那么中介效应本身就是研究的主题;如果是后者,那么研究中介效应只是为了更好地理解处理与结果之间的因果关系。

### 3. 如何理解貌似中介效应检验的做法

在文献中经常看到一类做法:先进行 $Y$ 对 $D$ 的回归,然后在回归中控制一个新的变量 $X$ ,并发现 $D$ 的系数依然显著,以此来论证 $D$ 对 $Y$ 的因果关系。用中介效应检验的术语来讲,这似乎意味着 $D$ 对 $Y$ 有“直接效应”。但这为何就能用来进行因果关系的论证呢?事实上,这种做法尽管看起来跟逐步法类似,但研究策略的出发点是截然不同的。再举两个例子。

Nunn and Wantchekon(2011)使用2005年非洲民意调查数据,检验非洲历史上的奴隶贸易造成的人际不信任是否延续至今。被解释变量是受访者报告的各种信任指标,反映受访者是否愿意信任他人,核心解释变量是受访者所在种族历史上的奴隶出口数量。结果显示,历史上奴隶出口越猖獗,如今人们就越不愿意信任他人。在“因果渠道检验”(Testing for Channels of Causality)中指出,奴隶贸易影响的可能是人们的内心行为规范(不愿意信任他人),也可能是外部环境(由于国家、制度和法律的衰败,他人不再值得信任)。他们想要强调的是前一种因果渠道,于是构造了能够反映后一种竞争性解释的指标并在回归中予以控制,比较控制前后核心解释变量的系数变化,来看奴隶贸易的效应是否被竞争性解释所吸收。例如,控制反映政府质量的变量(此时以对当地政府的信任作为被解释变量),奴隶贸易的系数变小了50%,这说明至少有一半的效应无法被外部环境的恶化所解释。又如,控制受访者所在地其他种族历史上受奴隶贸易的影响,从而间接控制周围他人的可信任程度(此时用族群间信任作为被解释变量),奴隶贸易的系数基本没有变化,说明奴隶贸易对今天族群间信任的效应无法被解释成奴隶贸易降低了他人的可信任程度,而只能解释为奴隶贸易使人更不愿意信任他人。

Rajan and Zingales(1998)认为,金融的主要作用之一是为生产性项目提供外部融资,金融市场的发展能够通过降低企业的外部融资成本从而刺激产业增长。核心解释变量是一国的金融发展水平和某一行业的外部融资依存度两个变量的交互项。这项研究在因果识别上的威胁是,金融发展水平可能反映的是能够促进产业增长的其他因素,而产业增长对这种促进因素的依赖可能与其对外部融资的依赖高度相关。他们提出的一个竞争性解释是,外部融资依存度高的行业可能人力资本需求也高,而人力资本发展水平可能跟金融发展水平相关。交互项系数可能反映的是,对人力资本需求越高的行业在人力资本发展水平越高的国家增长越快。于是,在回归中控制了平均受教育年限与外部融资依存度的交互项,发现核心解释变量(金融发展水平与外部融资依存度的交互项)系数并未发生改变,说明这个竞争性解释不足以威胁研究的主要结论。另一个竞争性解释是,外部融资依存度低的行业可能是成熟行业,金融发展水平低的国家可能是发展中国家。交互项系数可能反映的是,越成熟的行业由于技术转移越充分,因此在发展中国家增长越快。类似地,在回归中控制了人均收入水平与外部融资依存度的交互项,同样没有改变核心解释变量的估计结果。

在这类做法中,关于 $D$ 如何影响 $Y$ ,先有一个基准理论,然后提出一个竞争性理论,再构造能够刻画这种竞争性理论的控制变量 $X$ 放入基准回归,如果发现竞争性理论不能完全解释 $D$ 和 $Y$ 的相

<sup>①</sup> 提升能力的青少年经历可能和社交活动高度相关,再次说明不同的作用渠道难以严格分离。

关性(存在“直接效应”),则说明基准理论很可能是对的。这类做法和逐步法的区别在于, $D$ 不会影响 $X$ , $X$ 不是 $D$ 影响 $Y$ 的中介, $D$ 对 $Y$ 的因果“故事”和 $X$ 对 $Y$ 的因果“故事”是两个互相竞争的“故事”。 $Y$ 同时对 $D$ 和 $X$ 的回归有时被形象地称作“赛马”(Horse Race)。这种“貌似中介效应检验”的做法,实际上是用来强化因果关系论证的。

最后需要指出的是,本文是在约简主义方法论的框架下讨论中介效应分析。中介效应分析的传统在经济学中实际上由来已久,至少可以追溯到20世纪50年代用线性方程组描述的宏观经济模型,如Klein and Goldberger(1955)。经过“可信性革命”以后,在强调因果推断的研究语境中,经济学家并不热衷于心理学家所从事的那一套中介效应检验,这也就不奇怪了。但不可忽视的是,在经济学中同时还有另一支结构主义(Structural Approach)经验研究范式。在这一范式中,结构模型的构建往往需要依赖于对函数形式或不可观测变量的分布形式的较强假设,但与之伴随的优势是,研究者可以更好地运用经济学理论来认识现实数据的生成过程,严格地推导待估计的因果关系(而不仅仅是用理论来直观地解释估计到的因果关系),这也使得严谨的中介效应分析得以可能。特别地,来自随机实验的数据与结构主义方法论相结合,使得研究者可以部分地放松和检验结构模型的识别假设,一方面结构模型的可信度提高,另一方面来自随机实验的约简式因果关系的作用机理可以得到更深入地理解。这方面一个新近的例子是Heckman et al.(2013)。对此的全面讨论超出了本文的范畴。

## 六、正确认识调节效应分析

### 1. 调节效应分析与异质性分析的关系

调节效应分析和异质性分析这两者是一回事。最简单的理解:当调节变量 $M$ 是虚拟变量时,相当于把全样本分为 $M=0$ 和 $M=1$ 两个组,交互项 $M \times D$ 的系数就是分组进行的 $Y$ 对 $D$ 的回归中 $D$ 的系数的组间异质性。当 $M$ 是连续变量时,本质上并没有发生变化, $D$ 对 $Y$ 的因果效应受到 $M$ 的调节,也就是可以理解为, $D$ 对 $Y$ 的因果效应在高 $M$ 组和低 $M$ 组之间存在异质性。

如今研究者习惯做异质性分析,但是在研究设计阶段很少问为什么要做异质性分析。也许因果效应的异质性本身就是重要的。例如,在教育回报率的研究中,除了得到一个全样本的点估计之外,研究者可能还会关心女性的教育回报率是否显著地高于或者低于男性,由此得出不同的政策含义。

想要将因果效应的异质性本身作为文章的主要“卖点”,势必要求这种异质性的对比十分鲜明。比方说,因果效应在全样本中是显著的,但这种显著性只在其中一个子样本中继续存在,在另一个子样本中则不存在。例如,研究小额信贷( $D$ )对家庭财务状况( $Y$ )的影响,可能发现总体上两者呈现出反直觉的负相关——借了钱的家庭反而更容易陷入窘境;但对家庭按理财素养( $M$ )进行分组后发现,这个负面效应只在理财素养低的家庭中存在,在理财素养高的家庭中并不存在。这样的结果既在意料之外,又在情理之中。

但在一些文献中,研究者往往只是出于扩充文章篇幅的需要,在基准回归之外,出于某种“八股本能”,按地区、规模、所有制等进行一些异质性分析。反正这样做很安全,不管系数是否有差异,不管差异的方向是正是反,都有话可说。例如,某种效应如果在东部地区比较显著,可以解释为东部地区的市场化程度较高;如果在西部地区比较显著,则可以解释为西部地区受到国家优惠政策的倾斜更大。再比如,某种效应如果对非国有企业比较显著,可以解释为非国有企业的经济激励更强;如果对国有企业比较显著,则可以解释为国有企业的非经济激励更强。这种缺乏事前理论分析与预期,完全根据事后估计结果强行附会理论解释的做法,无异于数据挖掘。



一篇因果推断经验研究文章的重点永远是正确识别处理变量对结果变量的因果关系,因此,文章中的每一字每一句都应该为这一目标服务。而异质性分析更重要的作用正是通过分析因果关系的作用机制来强化因果关系论证。

2. 通过调节效应分析论证因果关系

仍然以 Rajan and Zingales(1998)为例展开讨论。在其引言部分明确指出:“要推进对金融与增长之间因果关系的理解(Make Progress on Causality),一种方式是关注金融发展影响经济增长的理论机制的细节。”“金融发展理应更好地帮助特别依赖外部融资的企业或行业的增长。这样的发现能够成为金融与增长因果关系辩论中的决定性证据(Smoking Gun)。”“通过寻找金融影响增长的一种具体机制的证据,能够为因果关系提供更有力的检验(A Stronger Test of Causality)。”如何理解这样的论断?下面来详细说明。

如果从统计上发现了  $D$  与  $Y$  的相关性,并且想要主张  $D$  是  $Y$  的原因,那么可以通过检验  $D$  影响  $Y$  的某个具体机制来对从  $D$  到  $Y$  的因果关系进行论证。论证的逻辑如下:①提出一个  $D$  影响  $Y$  的理论  $T$ 。根据这个理论, $D$  通过某个机制影响  $Y$ ,并且可以识别出这一机制在某些子总体中存在,在另一些子总体中不存在,令  $M = 1$  表示存在这一机制, $M = 0$  表示不存在这一机制。②在  $M = 1$  组,发现  $D$  与  $Y$  的相关性继续存在,而在  $M = 0$  组, $D$  与  $Y$  的相关性不复存在。③可能导致  $D$  与  $Y$  出现相关性的竞争性解释还包括  $Y$  影响  $D$  的反向因果理论  $R$ ,或者有混淆因素同时影响  $D$  和  $Y$  的遗漏变量理论  $C$ 。如果无法想象理论  $R$  或理论  $C$  发挥作用的机制在  $M = 1$  和  $M = 0$  组存在差异,则理论  $R$  或理论  $C$  很可能不成立。否则,应该在  $M = 0$  组也观察到  $D$  与  $Y$  的相关性。这样就完成了因果关系的强论证。这一逻辑的总结见表 1。

表 1 通过调节效应分析进行因果关系的强论证

$D$ 与 $Y$ 相关性的理论预期	$M = 1$	$M = 0$	对理论的判断
理论 $T$	存在	不存在	起作用
理论 $R$	同时存在或同时不存在		不起作用
理论 $C$	同时存在或同时不存在		不起作用
$D$ 与 $Y$ 相关性的数据检验	显著	不显著	

有时两组中  $D$  与  $Y$  的相关性都存在,但在  $M = 1$  组这种相关性更强,表现在  $Y$  对  $D$  的回归中  $D$  的系数估计绝对值在  $M = 1$  组更大,且组间差异在统计上显著。这时至少可以说, $D$  与  $Y$  的相关性不全是理论  $R$  或理论  $C$  所带来的,否则这种相关性应该在  $M = 1$  和  $M = 0$  组无差异。这样尽管没有证伪理论  $R$  或理论  $C$ ,但至少证实了理论  $T$ ,也在因果论证上迈出了一大步。这一弱论证的逻辑可以通过对表 1 稍作修改来总结,如表 2 所示。

表 2 通过调节效应分析进行因果关系的弱论证

$D$ 与 $Y$ 相关性的理论预期	$M = 1$	$M = 0$	对理论的判断
理论 $T$	存在	不存在	一定起作用
理论 $R$	同时存在或同时不存在		可能起作用
理论 $C$	同时存在或同时不存在		可能起作用
$D$ 与 $Y$ 相关性的数据检验	强	弱	

在 Rajan and Zingales (1998) 中,金融发展水平 ( $D$ ) 与经济增长 ( $Y$ ) 强相关,试图证明金融发展是经济增长的原因,并检验了金融发展通过缓解企业的外部融资约束来促进企业成长 (理论 T)。将行业分成两组<sup>①</sup>,一组是外部融资依存度较高的行业 ( $M=1$ ),另一组是外部融资依存度较低的行业 ( $M=0$ ),发现在对行业增长的回归中,金融发展水平与外部融资依存度的交互项显著,表明金融发展水平与行业增长之间的相关性在外部融资依存度不同的组间存在显著差异。金融与增长之间的相关性可能是因为增长影响金融,高增长引发了融资需求从而导致金融市场发展 (理论 R),也可能是因为某个混淆因素 (如节俭传统) 同时影响金融发展和经济增长 (理论 C),那么除非理论 R 和理论 C 在外部融资依存度不同的组间发挥作用的程度不同,否则就证明了理论 T。

再举一例。Brown (2011) 试图验证竞争是一种重要的激励机制。但考核相对绩效的锦标赛机制要想发挥作用,得有一个前提——竞争者的能力必须相对均衡,当存在能力超强的“超级明星”时,锦标赛机制反而可能出现负面效果。采用 1999—2006 年高尔夫球赛事中选手们的成绩数据,发现在“超级明星”——“老虎”伍兹参加的赛事中,其他选手的表现 (用总杆数衡量) 比在伍兹缺席的赛事中选手们的表现要更好。这意味着,与伍兹同场竞技时,其他选手受到了负向激励 (因为夺冠无望而不正常发挥)。由于伍兹参赛并不是一个随机事件,伍兹是否参赛与其他选手表现之间的负相关性存在竞争性解释:有可能因为伍兹参加的都是难度较高的赛事,而在高难度赛事中其他选手发挥相对较差是很自然的事,与激励无关。为了排除这种竞争性假说,进一步考察了伍兹参赛与否所带来的成绩反差在高水平选手还是低水平选手中体现得更明显。赛事难度的解释将预期到低水平选手的反差更大 (低水平选手更难以适应高难度赛事),负向激励的解释将预期到高水平选手的反差更大 (高水平选手更难以接受伍兹剥夺了他们夺冠的机会)。结果发现,高水平选手的表现反差大于低水平选手,这就证明了“超级明星效应” (尽管无法完全排除赛事难度的影响)。

Muralidharan and Prakash (2017) 使用三重差分方法评估了印度 Bihar 省为 14—15 岁女生免费提供自行车的政策是否提高了女生的入学率。比较了 Bihar 省受到政策影响的 14—15 岁女生入学率与未受到政策影响的 16—17 岁女生入学率的差异,减去同样年龄段男生入学率的差异,再与没有出台同一政策的临近省 Jharkhand 的同一双重差分相比较。但三重差分估计仍然可能受到政策实施时点其他不可观测因素的影响。为此考察这一政策影响入学率的机制——三重差分估计值是否随着离学校的距离而发生变化 (这本质上是四重差分),并发现两者存在倒 U 型关系,这与自行车降低通勤成本的理论预期相符 (家离学校过近,不需要自行车;家离学校过远,自行车帮助不大)。将这个结果作为文章最重要的结果,并在文中明确指出处理效应的异质性是一种检验识别假设的手段:“尽管三重差分的平行趋势没有被拒绝,但估计仍然有可能受到遗漏变量的潜在影响,如女生相对于男生的超额教育回报的增长趋势在不同省份之间存在差异,从而导致女生相对于男生的入学率差异在不同省份之间有所不同,因此有必要考察免费提供自行车政策对改善入学便利的作用机制。”

### 3. 好的调节变量

好的调节变量本身应该比较稳定,或者其变动是外生的,不受处理变量或结果变量的影响。Balli and Sorensen (2013) 讨论了这个问题。如果  $D$  影响  $M$ ,那么交互项  $M \times D$  可能捕捉的是  $D^2$  的效应,也就是说, $M$  对  $D$  与  $Y$  关系的“调节”实际上反映的可能是  $D$  对  $Y$  的非线性影响。但这不是最重要的理由。内生的调节变量相当于在双重差分研究中,处理组和控制组的构成一直在变化 (Compositional Change),并且导致这种变化的因素和  $Y$  相关 (隐藏在扰动项之中),这就很难说处理

<sup>①</sup> 在原文中  $M$  是连续变量,此处表述成虚拟变量更便于理解。

组和控制组的平行趋势假定还能成立。

在 Rajan and Zingales(1998)中,调节变量“外部融资依存度”的定义是企业资本性支出中的外部融资占比(再取行业中位数)。这个变量本身是资金市场上供求均衡的结果,高外部融资占比既可能反映行业需求面的技术特征(是否依赖外部融资),也可能反映金融市场发展水平的供给特征(外部资金是否充裕),因此不适合用一国一行业的实际外部融资占比作为调节变量。他们的做法是,考虑到美国的金融市场发展相对完善(资金供给具有充分弹性),美国企业的实际外部融资占比可能更多反映的是企业的资金需求,因此,可以用这个指标在美国各行业的取值作为其他国家相应行业的外部融资依存度指标(同时将美国从估计样本中剔除)。

图 10 反映了如果用各国各行业外部融资占比的实际数字作为调节变量将会产生的估计偏误。图中,圆圈内数字 1 和 2 为低外部融资依存行业,3 和 4 为高外部融资依存行业,细实线反映这两类行业在低金融发展水平国家和高金融发展水平国家间的增长差异,细虚线表示用低外部融资依存行业的实际差异所推测的高外部融资依存行业(假若没有从金融发展中受益)的反事实差异,因此,细实线和细虚线之间的距离就是金融影响增长的真实因果效应。如

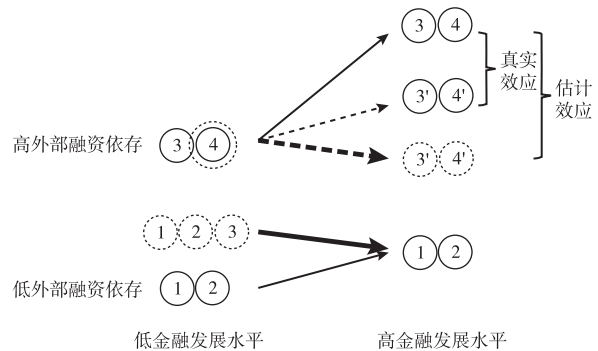


图 10 内生调节变量造成的估计偏误

果用各国各行业的实际外部融资占比作为分组依据,行业 3 本来是高外部融资依存行业,但由于在低金融发展水平国家中实际外部融资占比较低,被错误地归为低外部融资依存行业,此时低外部融资依存行业在两类国家间的增长差异就体现为粗实线,由此推测的高外部融资依存行业的反事实差异体现为粗虚线,这样就会高估金融对增长的因果效应。

朱家祥和张文睿(2021)认为,交互项在起始模型里就该出现,交互项系数统计显著是研究调节效应的起点,而不是终点。他们给出的理由是,如果先估计不包含交互项的基准模型,然后再引入调节变量和交互项,若交互项系数显著,就意味着在估计基准模型时,调节变量和交互项被留在了扰动项中,因此基准模型的估计是不一致的。如果调节变量本身是高度内生的,那么这个问题确实存在。但如果调节变量是外生的,在基准回归中遗漏它及其与处理变量的交互项,就不会造成估计偏误。此时,基准回归中处理变量的系数估计应该接近于交互项模型中的平均观测个体的边际效应。

#### 4. 作为因果识别第二种基本策略的调节效应分析

在因果推断研究中,研究者处理内生性的主要思路往往聚焦在寻找合适的控制变量和控制策略,即找到导致内生性的原因,然后正式地刻画、测量和控制它。调节效应分析则提供了另一种处理内生性的思路,即尝试挖掘因果模型的新的可验证含义——处理变量和结果变量之间更丰富的相关性,如果这种相关性是其他因果“故事”所不能解释的,那么即便此时内生性仍然存在,但至少证明研究者所感兴趣的因果关系是存在的,否则这种更丰富的相关性不会出现。因此,调节效应分析应该被看作因果识别的第二种基本策略的重要实现手段。

在金融促进增长的例子中,无法完全控制反向因果或第三方混淆因素(尽管可以控制不变的行业特征和国家特征,但仍然可能存在同时随行业和国家变化的遗漏变量),作者转而去挖掘更丰富的相关性:金融和增长的正相关性在外部融资依存度更高的行业是否更强?这个事实可以被金融通过缓解企业外部融资约束从而促进增长的“故事”所解释,但不能被其他“故事”合理解释,因果

论证的目的就达到了。同样地,在超级明星效应的例子中,无法完全控制“赛事难度”这一导致处理变量“伍兹是否参赛”存在内生性的因素(尽管可以最大限度地控制赛事级别、场地质量、奖金总额等),作者转而去挖掘更丰富的相关性:伍兹是否参赛导致的比赛杆数差异对高水平选手而言是否更大?这个事实可以被超级明星效应所解释,但不能被伍兹参加的都是高难度赛事所解释,因果论证的目的就达到了。

调节效应分析是一种很重要的因果论证手段,但在使用这种手段之前,首先要发展出一个说得通的理论:如果金融促进增长这个“故事”成立,那么就应该看到不同行业的效应大小不同(因为对金融的需求不同);如果超级明星效应这个“故事”成立,那么就应该看到不同选手的效应大小不同(因为激励强度不同)。然后再构建相应的交互项模型去验证这个理论。这就是因果推断理论先行(Theory Driven)的含义。也许可以先通过大量的尝试去发现稳定的相关关系,然后再试图给出理论解释,但这样的研究路径成功的机率不高,因为调节变量往往不会“躺”在数据集里等着研究者去发现,而需要研究者根据理论去构造。

## 七、国内应用现状与建议

### 1. 中介效应分析的应用现状

在经济学因果推断经验研究中使用中介效应检验是国内学术界一种独特的现象。其流行主要受到温忠麟等(2004)、温忠麟等(2005)、温忠麟和叶宝娟(2014)等研究的影响。不过,温忠麟等(2004)<sup>①</sup>明确指出,在中介效应检验中,因变量和自变量之间的关系不一定是因果关系,而可能只是相关关系。温忠麟和叶宝娟(2014)进一步说明,中介效应检验不能验证因果关系,中介效应模型中的因果关系都要有理有据,或者有某种学科理论支持,或者有文献做铺垫,或者有经验常识作为佐证。总之,因果链中的每一个关系都要在提出假设和建模之前得到支持。但是国内经济学经验研究中进行中介效应检验时,多数似乎并没有听从这样的忠告。

这里仅以国内某权威期刊近两年发表的文章为例加以说明,其中暴露的问题在目前国内研究中具有普遍性。在发表的全部文章中,涉及讨论因果关系作用渠道的有近30%,其中,近半数文章没有进行中介效应检验,只是分别估计了(1)式和(3)式<sup>②</sup>;超过半数的文章或多或少地进行了中介效应检验,具有如下特点:①有十余篇正式引用了温忠麟等(2004)、温忠麟和叶宝娟(2014)等相关的心理学文献。②绝大多数文章采用Baron and Kenny(1986)逐步法进行分析,其中有半数文章采用Sobel标准误或自助法对间接效应进行了统计检验;有1/4的文章关注了加入中介变量前后处理变量系数的变化,但没有对系数差异进行正式的统计检验;其余文章仅关注了系数估计的符号和统计显著性。③个别文章方法运用有误。例如,有的文章没有估计(3)式;有的文章在估计(1)式和(2)式时没有确保使用相同的估计样本,从而使得直接效应与间接效应之和不等于总效应;有的文章分别检验多个中介变量后,直接把间接效应显著的中介变量视为“主要中介”(尽管间接效应占总效应的比例很低)。还有一些不明所以的做法,例如,有的文章用加入中介变量之后处理变量系数估计显著性的下降(而非系数估计绝对值的下降)来论证中介效应的存在;有的文章通过(3)式

① 该文把Baron and Kenny(1986)方法引介到国内学术界。

② 其中部分文章还进行了Y对M的回归,但由于这些文章没有正式处理M潜在的内生性,因此这一结果只能被视为相关性证据,M对Y的因果效应仍然更多地依赖于理论直觉。

得到  $M$  的拟合值  $\hat{M}$ , 然后用  $Y$  对  $\hat{M}$  的回归来考察中介效应。④最重要的是, 所有文章都没有讨论中介变量在(2)式中可能存在的内生性问题。⑤有的文章甚至把中介效应检验视作稳健性检验的一种手段, 这意味着, 似乎只要“发现”了  $D$  对  $Y$  的间接效应, 从  $D$  到  $Y$  的因果链条就能自动变得更加令人信服, 这是大错特错的。中介效应分析的根本目的是考察  $D$  对  $Y$  的因果关系的作用渠道, 这是  $D$  对  $Y$  的因果关系研究的一个扩展。如果做对了中介效应分析, 自有其独立存在的价值, 但不能用来论证  $D$  对  $Y$  的因果关系。换言之, 在  $Y$  对  $D$  的回归中加入中介变量  $M$  后, 不论  $D$  的系数是否发生变化, 这一结果都无法使得  $D$  对  $Y$  的因果关系变得更加可信。

## 2. 中介效应分析的操作建议

关于中介效应分析, 本文提出以下操作建议: 一是停止使用中介效应的逐步法检验, 更不需要估计间接效应的大小并检验其统计显著性。把研究的重心重新聚焦到如何提高  $D$  对  $Y$  的因果关系的识别可信度。二是根据经济学理论, 提出一个或几个能够反映  $D$  对  $Y$  的作用渠道的中介变量  $M$ ,  $M$  对  $Y$  的影响应该是直接而显然的, 采用和第一条中同样的方法识别  $D$  对  $M$  的因果关系。三是尽量避免提出与  $Y$  的因果关系不明显、因果链条过长或者明显受到  $Y$  反向影响的中介变量。有时, 考察  $Y$  对  $M$  的回归也许是有益的, 但要记住这只是一条相关性证据。四是在绝大多数时候, 做好前两条就足够了。如果要考察  $D$  对  $Y$  的效应在多大程度上可以被  $M$  这一作用渠道所捕捉, 可以尝试在  $Y$  对  $D$  的回归中控制  $M$ , 但必须先弄清楚这种考察对理解  $D$  对  $Y$  的因果关系有何帮助, 并审慎解释回归结果。如有可能, 尽量论证这一结果受到  $M$  的潜在内生性的影响是有限的。当存在多个  $M$  时, 尤其要谨慎采用这种做法。

## 3. 调节效应分析的应用现状

调节效应分析的使用更为普遍, 除了少数例外, 调节效应分析几乎成为国内经济学经验研究的必要步骤。但是从对国内经济学领域一些权威期刊最近几年所刊发论文的阅读中可以看到, 这一研究实践尚存在如下主要问题: ①大多数文章没有把调节效应分析, 或者说机制检验, 提高到强化因果关系论证的高度。没有明确分析当前的研究方法仍受制于何种因果识别的威胁, 而特定的机制检验在何种意义上能够缓解这一威胁。②很多文章同时包含机制检验和异质性分析两部分, 此时异质性分析的目的并不明确, 很多时候只是为做而做, 而没有进一步分析这种异质性出现的原因, 即使稍有着墨, 也往往是用现象解释现象, 更谈不上讨论这种异质性如何服务于因果识别。甚至会出现两组异质性结果的理论解释互相冲突, 无法自洽。③有一些调节变量存在明显的内生性问题, 受到处理变量或结果变量的直接影响, 此时不但难以解释调节效应, 而且存在估计偏误。④个别文章使用同一个变量既作为调节变量又作为中介变量, 尽管理论上存在这种可能性, 但实际上很难分析清楚。理由很简单, 如果一个变量是调节变量, 那么分析调节效应时, 这个变量本身应控制在回归中; 但如果这个变量同时又是中介变量, 那么这种控制又会造成估计偏误。这往往反映出研究者混淆了作用渠道和作用机制之间的差异。⑤为数不少的文章采用分组回归的方式报告异质性分析结果, 但并没有对异质性进行正式的统计检验。

## 4. 调节效应分析的操作建议

关于调节效应分析, 本文提出如下操作建议: 一是将因果关系的作用机制检验视为因果识别的重要手段, 尽量正式地讨论其如何有助于强化对文章主题(从  $D$  到  $Y$  的因果关系)的论证。二是在研究设计部分详细阐述调节变量与调节效应的理论依据, 而不是等到报告实证结果时再附会解释。三是直观地展示调节效应, 讨论其数值大小在经济上的重要性。四是如果以处理效应的异质性本身作为研究目的, 明确说明这种异质性的经济意义——读者为什么要关心这种异质性。五是提高

统计规范性,对异质性进行正式的统计检验。六是将对作用机制和作用渠道的讨论进行严格区分,不宜安排在同一章节下,明确其不同的写作目的。

## 八、结语

本文是试图提升国内经济学因果推断经验研究规范性的一项努力。中介效应分析和调节效应分析作为社会科学工作者积极探索和深入理解因果关系的重要研究手段,本就是应这种规范性要求而生的,体现了科研工作者试图拨开数据“迷雾”、洞察真实世界的不懈追求。但是由于社会科学的因果问题特别复杂,研究实践的初衷和最终呈现效果之间难免存在一定的错位。中介效应分析的问题在于方法使用过度,研究者需要充分认识中介效应逐步法检验的局限性,小心从事因果识别;调节效应分析的问题在于方法发挥不足,研究者需要充分认识调节效应分析对于夯实因果识别的重要意义,大胆进行因果论证。如此则经验研究成果的科学性和可信性将进一步提高,其学术影响力和政策参考价值也将进一步提高。

本文的一般性意义在于,藉由对具体方法的讨论,重申定量社会科学因果推断的方法论立场。在因果推断中,较为艰巨的任务不在于统计方法的运用,而在于如何调用社会科学理论资源构建理解世界的方式。来自社会心理学家 Judd 和 Kenny 的告诫虽属老生常谈,但对向来标榜因果推断严谨性的经济学经验研究者来说,仍然值得一再强调:统计学是在因果模型正确的前提下开展工作,但因果模型很可能是错的,而统计学无法告诉我们错在哪里(Judd and Kenny,2010)。

## 〔参考文献〕

- 〔1〕温忠麟,侯杰泰,张雷. 调节效应与中介效应的比较和应用[J]. 心理学报,2005,(2): 268-274.
- 〔2〕温忠麟,叶宝娟. 中介效应分析:方法和模型进展[J]. 心理科学进展,2014,(5): 731-745.
- 〔3〕温忠麟,张雷,侯杰泰,刘红云. 中介效应检验程序及其应用[J]. 心理学报,2004,(5): 614-620.
- 〔4〕朱家祥,张文睿. 调节效应的陷阱[J]. 经济学(季刊),2021,(5): 1867-1876.
- 〔5〕Alesina, A., and E. Zhuravskaya. Segregation and the Quality of Government in a Cross Section of Countries[J]. American Economic Review, 2011, 101(5): 1872-1911.
- 〔6〕Angrist, J. D., and J. Pischke. Mostly Harmless Econometrics: An Empiricist's Companion[M]. New Jersey: Princeton University Press, 2009.
- 〔7〕Angrist, J. D., and J. Pischke. The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics[J]. Journal of Economic Perspectives, 2010, 24(2): 3-30.
- 〔8〕Balli, H. O., and B. E. Sørensen. Interaction Effects in Econometrics[J]. Empirical Economics, 2013, 45(1): 583-603.
- 〔9〕Baron, R. M., and D. A. Kenny. The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations[J]. Journal of Personality and Social Psychology, 1986, 51(6): 1173-1182.
- 〔10〕Bollen, K. A. Structural Equations with Latent Variables[M]. New Jersey: John Wiley & Sons, 1989.
- 〔11〕Brown, J. Quitters Never Win: The(Adverse) Incentive Effects of Competing with Superstars[J]. Journal of Political Economy, 2011, 119(5): 982-1013.
- 〔12〕Bullock, J. G., D. P. Green, and S. E. Ha. Yes, But What's the Mechanism? (Don't Expect an Easy Answer)[J]. Journal of Personality and Social Psychology, 2010, 98(4): 550-558.
- 〔13〕Dell, M. The Persistent Effects of Peru's Mining Mita[J]. Econometrica, 2010, 78(6): 1863-1903.
- 〔14〕Hayes, A. F. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach[M]. New York: The Guilford Press, 2022.



- [15] Hayes, A. F. Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium [J]. *Communication Monographs*, 2009, 76(4): 408–420.
- [16] Heckman, J., R. Pinto, and P. Savelyev. Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes [J]. *American Economic Review*, 2013, 103(6): 2052–2086.
- [17] Judd, C. M., and D. A. Kenny. Process Analysis: Estimating Mediation in Treatment Evaluations [J]. *Evaluation Review*, 1981a, 5(5): 602–619.
- [18] Judd, C. M., and D. A. Kenny. Estimating the Effects of Social Interventions [M]. New York: Cambridge University Press, 1981b.
- [19] Judd, C. M., and D. A. Kenny. Data Analysis in Social Psychology: Recent and Recurring Issues [A]. Fiske, S. T., D. T. Gilbert, and G. Lindzey. *Handbook of Social Psychology* (5th ed) [C]. New Jersey: John Wiley & Sons, 2010.
- [20] Kenny, D. A., D. Kashy, and N. Bolger. Data Analysis in Social Psychology [A]. Gilbert, D. T., S. T. Fiske, and G. Lindzey. *Handbook of Social Psychology* (4th ed) [C]. New York: McGraw-Hill, 1998.
- [21] Klein, L. R., and A. S. Goldberger. An Econometric Model of the United States, 1929–1952 [M]. Amsterdam: North-Holland Publishing Company, 1955.
- [22] Muralidharan, K., and N. Prakash. Cycling to School: Increasing Secondary School Enrollment for Girls in India [J]. *American Economic Journal: Applied Economics*, 2017, 9(3): 321–50.
- [23] Nunn, N., and L. Wantchekon. The Slave Trade and the Origins of Mistrust in Africa [J]. *American Economic Review*, 2011, 101(7): 3221–3252.
- [24] Persico, N., A. Postlewaite, and D. Silverman. The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height [J]. *Journal of Political Economy*, 2004, 112(5): 1019–1053.
- [25] Rajan R. G., and L. Zingales. Financial Dependence and Growth [J]. *American Economic Review*, 1998, 88(3): 559–586.
- [26] Robins, J. M., and S. Greenland. Identifiability and Exchangeability for Direct and Indirect Effects [J]. *Epidemiology*, 1992, 3(2): 143–155.
- [27] Shrout, P. E., and N. Bolger. Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations [J]. *Psychological Methods*, 2002, 7(4): 422–445.
- [28] Sobel, M. E. Asymptotic Confidence Intervals for Indirect Effects in Structural Equations Models [A]. Leinhardt S. *Sociological Methodology 1982* [C]. San Francisco: Jossey-Bass, 1982.

## Mediating Effects and Moderating Effects in Causal Inference

JIANG Ting

(School of Economics, Renmin University of China, Beijing 100872, China)

**Abstract:** Mediating-effects and moderating-effects analyses are common practices in empirical research of causal relationships, both with problems though. The main problem of mediating-effects analysis is the abuse of causal steps approach borrowed from social psychology. The main problem of moderating-effects analysis is a lack of focus on its role of justifying identification. The paper discusses in depth the bias stemmed from causal steps approach, how to do mediating-effects analysis properly, and how to use moderating-effects analysis to strengthen causal argument. Practical guidelines are provided so as to improve the quality of current research conduct.

**Keywords:** causal inference; mediating effect; moderating effect

**JEL Classification:** C18 C21 C52

[责任编辑:覃毅]