

## 社会科学中的因果关系及其分析方法

章奇\*

**【内容提要】** 本文讨论了因果关系的定义以及社会科学中如何分析因果关系及其必要性，介绍了社会科学中分析因果关系的几种主要的定性、定量分析方法，并且从理论上讨论了在使用这些方法时所需要关注的几个值得关注的理论问题。

**【关键词】** 社会科学、因果关系、定性分析、定量分析

## Causation and Causal Analysis in Social Sciences

Qi Zhang

**Abstract:** This article discusses what the definition of causation is and the tools of causal analysis. We in turn introduce some principal qualitative as well as quantitative methods that are widely used for the aim of detecting and measuring causal relationship in social sciences. We also discuss some theoretical issues we may encounter when we use these methods and tools to unravel causal analysis.

**Key Word:** Social Sciences    Causation    Qualitative & Quantitative Analysis

---

\*美国西北大学（Northwestern University）政治学系博士候选人，电子邮件：[qizhang2008@u.northwestern.edu](mailto:qizhang2008@u.northwestern.edu)。作者感谢与王勇、李志赟、谢韬、陈云松、单伟的讨论和建议，本文文责自负。

社会科学的主要任务和使命在于面对纷繁芜杂的现象世界，探索、厘清事件的真相，并寻找、鉴别隐藏在事件背后并导致现象发生的原因。其中，寻求真相受到各种因素的影响，且很多因素并不在社会科学研究者的掌握范围之内。但是，识别、鉴定现象背后的原因，却是社会科学学者责无旁贷的任务，且从复杂的现象当中寻找出其原因，并用简洁、严谨、合适的语言表达出来，也是判断他们工作价值的一个重要标准。不仅如此，找出现象之间的因果关系还具有一定的工具价值，因为一旦发现了真正的原因，往往意味着我们可以从干预、改变原因入手从而导致结果的变化，从而使干预、改造社会环境成为可能。

自改革开放以来，社会科学研究中一切以政治正确为标准，以意识形态为唯一指导的垄断局面不断被反思且益发显得不合时宜。与此同时，国外先进的研究方法和理念被不断介绍到国内，从而推动了我国社会科学的分工，使社会科学的发展有了一个质的飞跃。例如信息不对称理论、博弈论等理论，以及大量新的统计、计量工具，不仅在国内经济学研究中被广泛运用，对诸如政治学、社会学、历史学、法学等其他各种学科进步的促进作用也是显而易见的。

不过，尽管当前我国社会科学研究进步明显，但却鲜有人认真归纳总结如何分析现象之间的因果关系。许多研究者的工作往往仅停留在对现象之间关系的简单描述和统计，却忽略了所研究现象之间因果关系的性质，例如把简单的相关关系直接等同于因果关系，把虚假关系（Spurious Relation）当作真实关系，等等。另外，由于错误理解因果关系的性质，研究者还容易在模型设定、样本选择等方面犯下致命错误，从而使自己研究结果完全丧失有效性（Validity）。

因此，正确地分析因果关系，就构成提高我国社会科学学者研究质量的不可或缺的一环。而本文的目的，则是简明扼要地介绍一下目前国外学术界在分析因果关系方面所取得的进展和方法，以引起国内学术界的注意，并在这方面的的工作起到抛砖引玉的作用。<sup>1</sup>在文章的以下部分，笔者将首先介绍因果关系的种类，其次介绍几种分析因果关系的常用方法，最后讨论一下在分析因果关系时会经常碰到的一些问题以及相应的注意事项。

## 一、因果关系及其类别

一般而言，当我们谈到因果关系的时候，实际上是把不同变量放在一个时间维度里进行考察的。例如我们认为 X（因）导致了 Y（果），那么这包含两层意思：一是 X 的发生（变化）在时间上是先于 Y 的发生（变化）的，二是由于 X 的发生（变化），从而导致 Y 的发生（变化）。第一层意思相对明显，但第二层意思却相对复杂。例如，如果说 X 导致 Y，但同时 Z 导致 X，那么如何看待 Z 和 Y 之间的关系？如果除 X 导致 Y，而且 A 亦导致 Y，那么这三者之间又如何

---

<sup>1</sup> 原则上，本文所介绍和论述的因果关系分析方法也适用于自然科学。但社会条件的复杂性、主观性和不可逆性导致社会科学的因果分析不会是对自然科学分析方法的简单照搬和复制。换言之，两者的区别在于使用同一方法处理具体问题时所面临的约束条件不同，因此和自然科学相比，从社会科学中所得到的因果关系可能更多地受到时间和空间的限制，而不能被直接、简单地理解为抽象的、永恒不变的宇宙定律。

界定关系。再如，如果说在某环境下 X 导致 Y，那么是否意味着在另一个环境下我们会重复看到 X 导致 Y？反过来说，如果 X 导致 Y，那么是否我们消除 X 的影响，就可以避免 Y 的出现（发生）？等等。要对这些问题予以解答，我们需要对因果关系的种类进行分别。

### 1. 必要原因和充分原因

必要/充分原因有时也叫必要/充分条件。如果我们对干预现象的重复性发生（包括重复性促使或阻止、加快或减缓）感兴趣，那么就有必要搞清楚什么是必要原因和充分原因。

所谓必要原因，是指现象 Y 的发生（变化）一定以现象 X 的发生（变化）为前提，因此 X 被认为是 Y 的必要原因。换言之，如果 X 不存在，那么 Y 必定不会存在。例如，氧元素的存在是燃烧的必要条件，也就是说，一旦不存在氧，也就不存在燃烧。再看一句我们都熟悉的政治语言：“没有共产党就没有新中国”，它实际上也陈述的是“共产党”是“新中国”的必要条件，因而“没有”共产党就“没有”新中国。

必要原因的特点是我们能够从一个现象（果）的存在确定性地倒推回另一个现象（因）的存在。例如，如果我们能够知道取得博士学位是获得教授职称的必要条件，那么一旦我们观察到某人获得了教授职称，那么就肯定可以判断他在此之前获得了博士学位。同样地，如果货币过度发行是通货膨胀的必要条件，那么一旦我们发现出现通货膨胀现象，那么就肯定可以判断货币发行已经过量了，等等。如果我们的目的是阻止某些事件发生的话，那么必要原因则提供了一个有价值的思路：即阻止必要原因的发生。以货币发行与通货膨胀的关系为例，如果我们不想看到通货膨胀的发生，那么根据已知的必要原因，我们只需要把货币发行量控制住就可以达到目的了。

所谓充分原因，是指现象 X 的发生（变化）一定会导致现象 Y 的发生（变化）。例如在一定大气压条件下，气温下降到零摄氏度，水一定会变成冰，因此零摄氏度以下的温度是水变为冰的充分原因。而具有强烈的公民责任感的公民在选举时必定会到投票站去投票，因此强烈的公民责任感是公民外出投票的充分原因，等等。

需要注意到是，对于一个现象 Y，很可能有不同的充分原因（ $X_1$ 、 $X_2$ 、 $X_3$ ...）的存在导致这一现象的发生。例如心肌梗死、遭受致命车祸、中毒，等等，均可看作是导致一个人死亡的充分原因。贫穷、性别歧视、厌学，等等，均可看作是一个小女孩从学校退学的充分原因，等等。另外，对于社会现象而言，尽管有些原因可以既构成充分条件，也构成必要条件（例如货币发行过量之于通货膨胀），但在许多情况下，充分原因并不构成必要条件，反之亦然。例如，我们所提到的小女孩退学的例子，所有的原因都是充分原因，但却都不是必要条件。因此，即使我们看到 Y 的存在并且知道那些原因构成其发生的充分原因，但除非我们知道足够多的细节信息，我们无法像分析充分原因那样从结果直接倒推回原因。

必要/充分原因被看作是确定性的因果解释（Deterministic Explanation, or Determinism），因为要么原因是确定的，要么结果是确定的。这种解释和事件发生的概率可能性（Probabilism）相对立：事件的发生是完全非随机性的结果，如果方法得当，那么世界几乎是可以完全预测的。但实际上，有大量的因果关系并不是完全按照确定性的方式表现出来，而是以一种非确定性的，或类似回归模型的方式表现出来。

## 2. 非确定性原因

相对于确定性的必要/充分原因，非确定性原因通过两种方式表现出来：

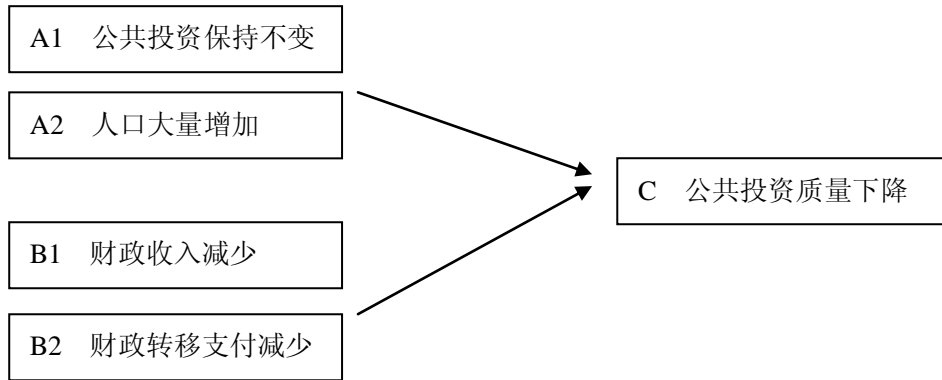
一种是解释变量除了可以被看作是完全的必要/充分原因外，也可以体现为在多大程度上被看作是近似必要/充分原因。例如，从完全性的必要/充分原因（Always necessary/sufficient）到近似的必要/充分原因（Usually necessary/sufficient）。Tomas Ertman (1997)是一个较好的例子。在其名著《利维坦的诞生：中世纪和早期现代欧洲的国家构建与政治体制》（The Birth of Leviathan: Building States and Regimes in Medieval and Early Modern Europe）中，Ertman 试图解释为什么在中世纪结束前后欧洲国家会形成不同的国家体制。他首先比较了以前的各种理论，发现如果用 14 个欧洲国家的例子来验证这些理论，那么在这 14 个案例中，至少有 4 个案例是和这些理论的预计相矛盾的。而他自己所提出的理论解释则和 14 个案例中的 12 个相吻合，只是在解释瑞典和丹麦时发生了错误。因为对这两个案例的解释失败，Ertman 无法将他的解释变量看作是完全的充分原因，而只能将其看作是近似的充分原因，但却比他所比较的其他理论更接近于充分原因（Mahoney 2003）。

另一种是各个解释变量都可以对结果（被解释变量）具有一定的作用，但都无法被独立看作是完全的充分或必要条件。例如，良好的地理条件、较高的储蓄率、较高的人力资本，等等，都对经济增长具有作用，但都无法被看作是经济增长的完全的充分或必要条件。而醉酒驾驶和车祸之间的关系（Liberson 1991）也与此类似。在回归分析中，这种关系通常被表述成：在其他条件不变的情况下，X 的  $\Delta x$  变化导致 Y 发生的概率发生  $\Delta y$  的变化。

## 3. 组合式非唯一充分原因（INUS）

INUS 为英文“Each is insufficient [I] but necessary [N] part of a condition which is itself unnecessary [U] but exclusively sufficient [S] for the effect”的缩写，意为对一组可以用来解释结果的解释原因而言，所有这些因素可以划归为若干子原因组，其中每一个子原因都可以构成导致结果的充分原因，而构成这些子原因的因素本身，对其所在的子原因而言都可以看作是构成它的必要而不是充分原因。图 1 给出了某地区公共投资质量下降的原因构成图，可以看作是一个典型的 INUS 例子。

图 1 INUS 示意图



在图 1 中，A1 和 A2 的组合，以及 B1 和 B2 的组合，分别构成结果（C）的（非唯一）充分原因。而 A1 和 A2 则分别是其自身组合的必要而非充分原因，B1 和 B2 则分别是其自身组合的必要而非充分原因。因此，INUS 实际上可以看作是一个多重路径的因果关系，一个结果可以通过不同的原因产生，而这些不同的原因又是各个因素的组合。这种关系可以通过如下语言表示：

$Y = (X_1 * X_2 * \dots) + (Z_1 * Z_2 * \dots)$ ，其中\*和+分别表示“并且”与“或者”的逻辑运算符号。

INUS 可以使我们对确定性原因（充分原因）和非确定性的原因进行进一步思考。从本质上讲，非确定性原因之所以存在，是因为人类认知能力的有限性和信息的不充分所致。在认知能力和信息充分的条件下，我们完全可以识别确定性原因。换言之，给定考察对象范围，即使一个单个的因素无法构成确定性的必要/充分原因，但它和其它因素的组合却完全可以构成导致结果的充分原因。

比如，在 Liberson (1992) 醉酒驾驶和车祸的例子中，醉酒驾驶并不能构成车祸的充分原因，但如果是在夜间行驶且亮度较低的情况下，三者的结合却成为出车祸的充分条件。同样，实行固定汇率制度并不是一个国家是否发生金融危机的充分原因，但如果一个实行固定汇率制度的经济发生较严重的通货膨胀（或贸易赤字），且不存在资本控制的条件下，那么这三者的组合却是发生金融危机的充分原因。简单地说，考虑 INUS 的存在，随着我们知识的积累，认知能力的上升，以及对细节信息掌握程度的提高，我们完全有可能找到导致 Y 发生的充分原因。

#### 4. 直接（Proximate）和间接（Remote）原因

直接原因是不经过其他因素而直接会导致结果发生的原因<sup>2</sup>，而间接原因是直接原因的原因。换言之，如果 X 是 Y 的直接原因，而 Z 是导致 X 的原因，那么 Z 就是 Y 的间接原因。例如，对某产业投资的决策而言，对该产业增长前景的信心程度是决策的直接原因，而该国的法律对产权的保护则是间接原因，因为它通过影响投资者对该产业的信心来影响投资决策。

<sup>2</sup> 如果有多重因果路径的话，那么就会存在多个直接原因的可能性，但其中只会有一个最终导致结果发生，该原因被称为触发（Precipitating）原因。

间接原因一般在分析历史因素时会起到很大的作用。例如在历史学上有个有趣的命题，即中国长城的修建导致罗马帝国的衰亡。其因果链条如下：长城的存在阻挡了野蛮人向东方挺进，迫使他们向西方发展。这些向西汹涌而去的野蛮人最终冲破了罗马帝国的防御，导致后者衰亡。这里，来自野蛮人的进攻是罗马帝国崩溃的直接原因，而中国长城的修建则是间接原因。

Acemoglu 等人 (Acemoglu et al 2001 2002) 提供了另一个直接和间接原因的例子。他们的研究指出，前欧洲殖民地国家战后的经济增长绩效取决于他们的制度质量，而制度质量又取决于 1492 年当欧洲殖民者开始到达殖民地时针对殖民地所作出的制度安排。这些殖民制度在战后殖民地获得独立后因为路径依赖的关系而继续发挥影响，从而影响到现时制度的质量，并最终影响经济增长。

在多大程度上应该找寻间接原因取决于所研究问题的性质以及间接原因对最终结果的作用。比如保险公司调查一栋办公楼起火的原因，直接原因是客户蓄谋放火谋取保险金，间接原因是因为某处农作物收成大幅下降，导致其期货投资亏损。这里尽管间接原因作用明显，但保险公司是不会对提高农作物收成产生任何兴趣的。而对于主管一国经济的决策者而言，资本和劳动投入的增长是经济增长的直接原因，法治环境的改善、自然环境的提升、政府职业培训投资力度的加大，等等可以看作是间接原因。这里决策者不仅关心间接原因，而且会努力找出所有有影响的间接原因。

另外，弄清楚直接和间接原因的关系在一定条件下对定量分析中的模型构建也是非常重要的。例如在回归分析中经常使用的工具变量法 (Instrument Method)，实际上就是通过寻找满足一定条件的间接原因作为直接原因的工具变量，以此来判断因果关系和衡量其大小。并且，这种间接原因作用于直接原因的路径一定需要满足唯一性 (即只能通过直接原因这一条路径影响结果)，否则工具变量法将不再有效。

## 5. 双向因果关系

除了以上所介绍的单向因果关系外，现实中还存在大量的双向因果关系。在一定时期、条件下，X 是 Y 的原因。随着时间、条件的改变，Y 成为 X 的原因。例如经济政策和经济危机的关系、经济增长和收入分配的关系、工资上升和物价上涨的关系，等等。不过这些双向因果关系，在很大程度上是建立在多样本基础之上，是对连续性、重复性现象的一般化。如果把注意力集中到单独的事件，那么单向的因果关系会更有助于我们对现象关系的理解。

## 二、分析因果关系的方法

分析因果关系的方法有很多，按其所使用的分析工具，一般可以分为两类：定量和定性方法。以下我们将主要介绍几种在社会科学研究中被广泛使用的定量和定性方法。

### 1. 定量方法 1—相关分析或回归分析法 (Regression Analysis, RA)

在 RA 当中，研究者一般在对数据作出一定假设的前提下（Gujarati 2003），将因变量（Y）对自变量（X1、X2...）进行回归，从而得到他们之间的数量关系。一般采取如下形式：

$$Y = \alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \dots + \beta_k \cdot X_k + \xi \quad (1)$$

根据这种方法，研究者可以在多样本观测的基础上得到自变量（解释变量） $X_i$ 对因变量（被解释变量）Y的平均影响。换言之，对单个样本观测（ $i$ ）而言，研究者定义因果效应采取如下形式：

$$\text{因果效应} = y_i^T - y_i^C \quad (2)$$

其中，上标 T 表示刺激（Treatment），C 表示控制（Control）。方程（2）说明单个观测到因果效应表现为被解释变量在解释变量变化前后的差异，这被称作是观测  $i$ “被实现的因果效应 (Realized Causal Effect)”（King, Keohane, and Verba 1994, 78-9）或者“单观测因果效应 (Individual Causal Effect)”（David 2000）。在实际中，很多情况下我们常常只能观测到样本接受刺激前或者刺激后的实际值，因此这种方法又被看作是“反事实的 (Counterfactual)”分析方法。在多样本观测到情况下，因果效应不再表现为单观测因果效应，而是表现为不同观测之间单观测因果效应的平均水平，这一效应通过方程（3）表现出来：

$$\text{因果效应} = \mu_i^T - \mu_i^C \quad (3)$$

（3）又被称为平均因果效应（Average Causal Effect, or Mean Causal Effect），其中  $\mu$  表示众样本接受刺激和控制状态下的平均效应。这样，在单个观察条件下无法被直接观测到的单观测因果效应就被转换为多观测条件下的平均因果效应。

可以看出，回归分析法是通过衡量变量 Y 在 X 变量接受刺激和接受控制之间的状态变化来衡量变量 X 所导致的 Y 的变化，因而这种方法也被看作是“因果关系效应法 (Effect-of-Causal)””。而这种衡量平均效应的方法并不适合用来直接检验确定性的必要/充分原因。在回归方法下，X 对 Y 的因果关系一般表示为：在其他条件不变的情况下，X 的一定变化所导致的 Y 的相应变化。

## 2. 定量分析法 2—实验分析法 (Experimental Design, or Experimental Analysis)

实验分析法是在上世纪 60、70 年代开始逐步被应用在社会科学研究中，近年来更是被广泛应用到经济学、政治学和心理学等领域。在公共物品投资、政治参与、收入分配、激励机制、行为主体的心理感受、认知能力对行为影响等各个方面，实验方法所带来的冲击和效果都是非常显著的。实验方法的基本思路是，通过实验设计者人直接控制、操纵实验参与者所面临的约束和激励条件（解释变量 X），并观察他们在接受刺激前、后的状态（解释变量 Y）的变化，从而得到 X 对 Y 的因果效应。由于在实验中，X 的变化是完全外生的，因而回归分析中所经常困扰研究者的内生性问题（Endogeneity Problem）在很大程度上得到了控

制。

研究者指出，一个高质量的实验设计需要考虑实验的内部（Internal Validity）有效性和外部有效性（External Validity）问题（Campbell and Stanley 1966）。所谓内部有效性问题，主要指因为实验设计本身存在的贻误，导致实验结论的有效性受到怀疑，这主要包括以下几个方面：

（1）历史问题（History）。如果实验接受刺激前和接受刺激后这两种状态之间相隔超出一定时间（例如数天、数月，等等），那么即使两种状态出现差异，这种差异也不一定能够归结为因为接受刺激所带来的变化，而有可能是在此期间所发生的其他事件导致接受刺激前后状态出现变化。比如研究者想观察实验接受者观看黑人文化纪录片对种族政策态度的作用，但在播放纪录片期间（比如持续了一周），一些反映黑人受歧视的新闻报道正在电视里频繁播出且实验接受者观看了这些新闻，那么这几乎可以肯定会对实验接受者的态度产生影响。

（2）成长问题（Maturation）。这一问题和实验过程中，时间流逝所自然带来的问题相关。比如他们因为实验时间拖得较长而显得不耐烦、疲倦，甚至因为时间足够长而对问题突然有了新的看法，那么实验中所观测到的差异也不能简单地归结为接受刺激所带来的变化。和历史问题相比，成长问题看重的是一个连续的过程，而前者则注意单独的事件的影响。

（3）测试问题（Testing）。一方面，随着实验接受者对实验程序和规则的熟悉，因此实验的次数越多，他们对接受刺激的刺激会越来越熟练。另外，聪明的实验接受者可能在实验中猜测出实验设计者的意图，因而会有意以一种使实验设计者满意的方式进行反应。这些因素会导致因果关系的推导出现偏差。

（4）工具问题（Instrumentation）。如果接受刺激前后测量状态的方法或标准出现变化，而这种变化和实验本身无关，那么接受状态刺激前后的差异就不能直接用来进行因果关系判断。

（5）回归问题（Regression）。如果接受刺激前端状态本身就是一种极端情况，那么无论是否接受刺激，其状态都会向正常方向回归。同样地，这也会给因果关系的判断带来偏差。

和内部有效性不同，外部有效性问题和实验本身用来衡量因果效应的设计无关，但却会影响从实验中所得出结论的一般化（Generalizability），即实验结论的有效范围是有限的，很可能只局限于实验所覆盖的样本之内。有些影响内部性有效性的问题，例如从成长问题可以看出，适用于少年人的实验结论可能无法推广到成年人。从逻辑上说，外部有效性要求从现有样本中得到的结论被外推到其他样本中去，而后者并不一定和实验样本具有同样的特性。这样，外部有效性问题可以被减弱，但永远无法彻底解决。

为了尽可能地增加实验设计的外部性和内部有效性，研究者指出在一个较理想的实验中，刺激（Treatment）的实施需要是随机的，和 Y 的初始状态不存在相关关系。实验接受者的接受刺激前后的状态应该是互不相关的。样本的选取应该遵循随机原则，且样本在对比组（Treatment Group）和控制组（Control Group）



之间的分配也应该是随机的。<sup>3</sup>图 2 给出了一个在实验中经常被用到的设计程序，它简单地说明了在设置好了对比组和控制组织后，实验是怎样被贯彻的。

图 2 简明实验设计程序

R O1 X O2 (对比组)  
R O3 O4 (控制组)

其中 R 代表对样本在实验各组之间的随机配置；X 代表对实验设计者所施加的刺激措施，它代表了解释变量 X，它对实验发生刺激前后状态差异的作用正是实验设计者所感兴趣的；Os 代表实验所处的状态，它可以由某种方式或指标衡量出来。每一行从左到右表示过程或时间顺序，其中处于 X 左边的被看作是接受刺激前 (Pretest) 时期，处于 X 右边的被看作是接受刺激后 (Posttest) 时期。通过对比组接受刺激前后状态的差异的对比，以及这一差异和控制组相应时期状态差异之间的对比，实验者可以对因果关系及其效应进行判断。

不过，社会科学的实验并不全部是在实验室里面完成的。根据实验发生的条件、环境和作用对象等实验构成因素的不同，以下两类实验也是社会科学研究者经常使用的方法：自然实验 (Natural Experiment) 和实地实验 (Field Experiment)。前者是指那些具有一切实验室实验特征的一整套过程，但其中的控制组和对比组，以及刺激的发生，都是在完全的自然环境中形成，而不是人为实验设计的结果。例如二战后南北朝鲜的形成及其经济发展实践被经济学家看成是一个对比经济制度绩效的自然实验。二战结束后，南北双方在美苏的支持下各自选择了迥然不同的经济体制。前者保留了市场机制实行市场经济，而后者则仿效斯大林模式实行计划经济。而在此之前，南北双方享有相同的文化和传统，地理位置和环境、自然资源禀赋。实际上，仅从工业化水平上来看，北朝鲜还略占优势，因为在日本殖民统治时期，大部分工厂都建立在北方而不是南方。这样一来，50 年前的领土分离，南北双方选择不同的制度就为判断制度如何影响其后经济发展的绩效提供了一个绝佳的自然实验案例。

所谓实地实验，是指在自然环境中（而不是实验室里）进行贯彻实验设计，实验对象尽量从实际社会中选取、实验所涉及的一切手段和刺激也尽量满足和实际相符合的原则。因为田野实验和实际环境结合较为紧密，因而该方法被认为是更好地解决了实验外部有效性的问题。另外，即使从数据质量上讲，实地实验也存在一些优点。比如传统的实验室试验里，参加对象主要是大学里的学生。他们的人口统计特征 (Demographic Characteristics)，包括年龄、教育背景，等等均受到很大限制。如果是实地实验，由于样本从当地人口中随机抽选，则人口统计特征的限制就被大大缓解了。这意味着样本的差异性得到了提高，对数据的统计

<sup>3</sup> 无法保证随机原则的实验，一般被称为准实验设计 (Quasi-experimental Design)。例如对历史现象进行分析时，作为被分析对象的历史过程可以被看作是一个准实验设计过程。

分析和结果在更大范围内的适用性都有很大的好处。不过，也有研究者指出和原始意义上的实验室实验相比，两者之间的区别只是程度上的，并不存在明显的优劣之分（Harrison and List 2004）。

### 3. 定性方法 1—米尔分析法（Mill's Methods）<sup>4</sup>

米尔分析法被广泛用来分析在小样本条件下确定性的必要/充分原因是否存在。这一方法主要包括：

#### （1）一致分析法（Method of Agreement, MOA）

MOA 可以用来确定地排除那些不构成必要原因的因素。MOA 的核心含义用表 1 可以直观简洁地表示。其中第二列中的字母分别代表潜在的解释变量，第三列代表被解释变量的发生。

表 1 MOA 法示例

案例	解释变量	被解释现象
1	A B C	S
2	A E F	S
3	A H G	S

在第一行中，A、B、C 的事前存在和现象 S 的随后发生联系在一起。在第二行中，A、E、F 的事前存在和现象 S 的随后发生联系在一起。在第三行中，A、H、G 的事前存在和现象 S 的随后发生联系在一起。A 的存在总是和 S 的发生相联系，而 B、C、E、F、H、G 消失则不影响 S 的发生。因此我们可以认为 A 是 S 的必要原因（不过这一结论却有可能因为新案例的发现而改变），而 B、C、E、F、H、G 是 S 的必要原因的可能性则可以被确定排除。

#### （2）差异分析法（Method of Difference, MOD）

MOD 可以用来确定地排除那些不构成充分原因的因素。MOD 的核心含义用表 2 可以直观简洁地表示。

表 2 MOD 法示例

案例	解释变量	被解释现象
1	A B C D	S
2	- B C D	-

在第一行中，A、B、C、D 的事前存在和现象 S 的随后发生联系在一起。在第二行中，B、C、D 的事前存在以及 A 的消失和现象 S 的随后消失联系在一起。因此我们可以认为 A 是 S 发生的充分原因（不过这一结论却有可能因为新案例的发现而改变），而 B、C、D 是 S 的充分原因的可能性则可以被确定排除。

在 MOA 和 MOD 方法的基础上，米尔分析法还有很多变体，例如米尔差异-一致联合法，残差法、协同变化法等等。

米尔分析法不仅可以用来分析单个变量是否构成必要/充分原因，还可以用来分析多重路径的因果关系，例如 INUS 原因是否存在。例如考察如下导致一所

<sup>4</sup> 米尔方法因其发明者、英国哲学家 John Stuart Mill（1806-1873）而得名。

房屋发生火灾的 INUS 原因（图 3）：

图 3 火灾发生的 INUS 原因

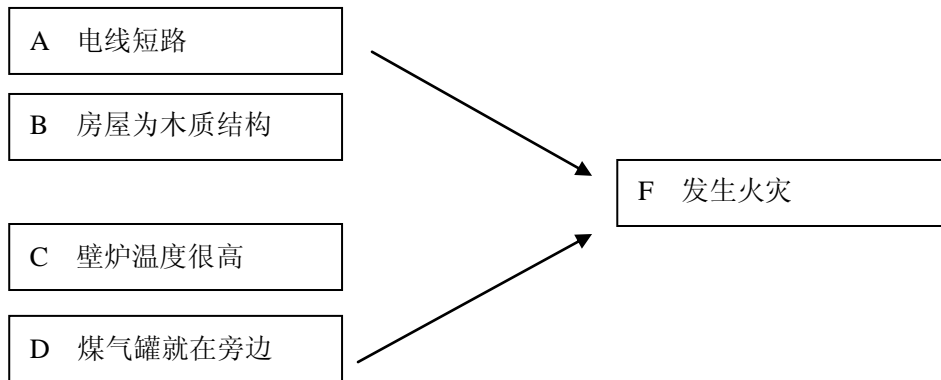


图 3 表明，A 和 B 的组合构成 F（火灾）发生的充分原因，C 和 D 的组合同样构成 F 发生的另一个充分原因。表 3 给出了利用米尔差异—一致联合法来反映这种关系：

表 3 差异—一致联合法与 INUS 原因

A	B	C	D	F
1	1	0	0	1
1	0	0	1	0
1	0	1	1	1
0	1	0	0	0
0	0	0	0	0
0	0	1	1	1
1	1	1	1	1

注：数字 1 代表该因素发生，0 代表没有发生。

从表 3 可以看出，A、B、C、D 中没有任何一个单独因素构成现象 F 发生的必要或充分原因。不过，A 和 B 的组合，以及 C 和 D 的组合构成 F 发生的 INUS 原因。

#### 4. 定性分析法 2—过程追踪法（Process Tracing）

过程追踪法一般被用于对一个或几个案例进行因果分析。但这种分析不是对某一试点特殊状态的分析，而是对一个连续性过程进行解剖，试图通过对历史发展轨迹前后顺序的解读来找出因果关系。不仅如此，这一方法还试图明确因果作用机制是怎样具体运作的，换言之，它希望回答 X 究竟通过什么样的具体作用方式和渠道导致了 Y（ $X \rightarrow Z \rightarrow Y$ ）。比如，一种假说认为资本主义的发展会导致民主制度最终出现。其作用机制为资本主义发展会壮大工人阶级以及中产阶级的力量，而工人阶级是坚决拥护民主的力量。资本主义的发展还会削弱大土地拥有者的力量，而这一阶级是坚决反对民主的。如果工人阶级在斗争中学会如何组织自己，并寻找到自己的盟友与之联合，那么民主就会出现。这样，资本主义的发展会导致拥护和反对民主的两大力量之间的相对强弱，从而成为民主制度的建立

创造必要条件。如果用过程分析法验证这一假说，就应该从民主制度建立之前，工业化伊始之时开始展开分析。观察是否资本主义的发展导致了工人阶级力量的壮大，是否工人阶级自身的斗争并与其他阶级之间的联合迫使大土地所有者阶级放弃政治垄断，从而实现民主。因此，这一过程实际上说明资本主义（X）是通过阶级力量相对变化（Z）这一机制起作用的。而对于那些没有出现阶级力量相对变化的案例，资本主义的发展不会导致民主制度的建立（Y）。任何一个环节（X→Z 或 Z→Y）如果没有被历史过程所证明，都会对原假说的因果关系予以了证伪。这种要求所有环节都必须没有被证伪的方法不仅丰富了我们因果关系作用机制的认识，而且也减小了 X 和 Y 之间出现的相关关系是虚假关系的可能性。

以上对最常见、常用的方法进行了概括。社会科学中用来分析因果关系的方法还有很多，比如定性方法还有模式匹配法（Pattern Matching）、因果叙述法（Causal Narrative）、Fuzzy-set 法，等等。对定量分析法而言，无论是回归分析法还是实验分析法，都包括极其庞大、复杂的分支方法，例如时间序列分析法、路径分析法、工具变量法，等等。本文不再也不可能一一列举。

### 三、一些相关的讨论和分析

研究者在分析具体问题，常常会思考如下问题：究竟如何从纷繁芜杂的现象之间发掘出因果原因？应该选取什么样的（定性或定量）方法来证明假说中的因果关系？如何选取样本？在前两节内容的基础上，本部分将对这些问题进行初步的分析和解答。

#### 1. 理论假说和因果分析方法：孰先孰后？

许多人认为，如果可以用以上所介绍的因果分析方法来证明或者证伪现象之间因果关系，那么能不能用这些方法来直接寻找理论，提出理论呢？一个最简单的例子就是，如果我们看到两个变量之间存在联系，那么是否意味着我们就可以从中得出一个具有因果原因意义上的理论以对此进行解释呢？回答是：几乎不能。

任何有价值的理论，都并非是对现象之间简单关系的刻画和描述，而是建立对现象之间细致观察和深刻理解的基础之上的。例如火烧皮肤会造成疼痛感、喝水可以解渴，自然灾害导致作物歉收，等等。可以认为这些陈述都反映出了一定的因果关系，但没有人认为他们具有多少理论上的价值。相反，理论本身是对重复性现象的高度抽象和概括，是对现象的重新判断、加工、组合与取舍。一个理论的提出，不但意味着一种因果关系被建立起来，而且代表着对现象本身含义的和与其他现象之间关系以及相对重要性的重新认识。而这些认识和知识不可能通过因果分析方法本身来得到。因此，理论是先于因果分析方法的，因果分析方法可以被用来检验理论，但我们无法通过后者来得到前者。

例如对于一个理论假说：一个国家民族构成越复杂，政治局面越不稳定。研究者考察了许多案例，在案例 1 中，某国的民族构成是存在 N1 个民族。在案例 2 中，某国的民族构成是存在 N2 个民族， $N1 < N2$ ，但后者的政治稳定性却高于前者。因此研究者认为，民族构成和政治稳定性具有反向的因果关系，从而否定了原理论假说。但是，研究者在这里实际上对民族构成的认识出现了偏误，如果他用是否具有信仰不同甚至敌对宗教的少数民族来分析民族构成的话，那么民族构成和政治稳定性之间的关系就符合原理论假说。但是，这种对民族构成分析方法的知识却并不能由分析因果关系的方法中得到，而只能由研究者对现象的深刻观察和高度抽象、归纳中得到。换言之，因果分析方法本身并不能保证我们得到正确的理论认识。

另外，即使我们通过一定的因果分析得出现象之间的因果关系，但如果我们考察的现象出现了关键性的遗漏，那么这种因果关系实际上仍然是错误的认识。例如，研究者试图考察公民对领导人是否满意以及国外是否提供经济援助的因果关系。考察的 Y1 是是否对领导人（A）满意（满意=1，不满意=0），Y2 是是否获得国外援助（得到=1，没得到=0）。而解释变量包括 A 的个人形象主观评价的好坏（X1，好=1 或不好=0）以及国外领导人对 A 的评价（X2，高=1 或不高=0）。运用 MOD 方法，得到的结果如下（表 4）：

表 4

案例	解释变量	被解释现象
1	X1=1, X2=1	Y1=1, Y2=1
2	X1=0, X2=1	Y1=0, Y2=1

根据 MOD 方法，我们可以认为 X1 是对领导人感到满意的成分原因。但实际上，研究者遗漏了另一个真实的原因——领导人实施了他在竞选中所承诺的政策（X3）。实际上，正确的理论应该是 X3 和 X1 共同起作用，导致公民对领导人的态度。但是，是否因该包括 X3 这样的知识并不能由具体的因果分析方法中得到。

另外，也是我们经常遇到的一个问题，就是即使我们发现两个变量之间具有很强的相关关系，在错误认识的条件下，也很难得到他们之间是否是因果关系的结论。例如，一个理论假说是是否能上好的学校和日后事业是否成功之间存在着因果关系。如果研究者的分析方法是找到一批包含了事业成功或不成功的职业人士，然后把他们之前所受教育经历进行对比，发现的确原来上好学校的学生现在的事业更成功，因而得出原有假说成立的结论。但实际上他的这个关于因果关系的判断却极可能是错误的。实际上，真正的原因是学生的家庭出身是否富裕，出身富裕的家庭使学生能够上好学校，接受好的教育。出身富裕也使他们在今后能找到更好的工作，搭建更好的交际网络，从而使他们事业更为成功。在这里，家庭出身是否富裕才是原因，而是否上好学校和事业是否成功才是结果。由于错误地认识了这三者之间的关系，研究者所得到的因果关系的结论也是错误的，尽管他所使用的验证方法是无错误的。但是否应该考虑进学生家庭出身这一因素，并不能由分析方法本身来提供和保证。

总之，提出一个反映因果关系的理论，并不依赖于用来分析因果关系究竟是否存在的方法是否完美和正确。后者取决于研究者的专业训练程度和经验，而前者除了这两条因素外，还取决于个人的研究兴趣，以及观察、抽象、和综合能力。

## 2. 定性和定量模型的选择

对此并不存在一个唯一的判断标准和回答。一般而言，究竟选择定性还是定量模型，取决于一系列的因素，包括：

(1) 研究的目的。前文已经说过，定量分析方法衡量的是解释变量  $X$  对被解释变量  $Y$  的平均作用效应。相对而言，定量分析方法考察的是对结果的具体原因解释，它很少涉及原因  $X$  具体程度的大小，更多地采用二分法（例如高一低、多—少、有—没有、好—坏，等等）来描述变量。因此，如果研究者不仅希望判断出是否存在因果关系，还希望得到有关作用平均效果的知识，那么就应该采用定量分析方法，比如对货币发行量和通货膨胀之间因果关系的分析。如果研究者更多的是希望找出变量之间的因果关系，且希望弄清楚“因”于“果”的具体作用机制，那么他们就应该采用定性分析。比如，社会资本和经济发展之间的因果关系。研究者当然可以通过定量分析法判断前者对后者的平均作用大小，但同样有价值的是社会资本通过什么样的具体渠道影响经济发展。比如，通过增强社会参与、增加公民之间的相互信任和合作，等等。对这些具体作用机制的分析，定性分析方法更具优势。

例如，从理论上讲，民主和经济增长之间存在着双向因果关系。如果研究者感兴趣的是讨论民主或经济增长究竟在什么样的具体条件下，通过什么样的具体渠道来影响对方，而不是平均而言，其中一个变量对另一个变量起到多大作用，那么定性方法应该是更好的选择。但如果研究者的目的并不在于界定具体而言两者如何互相作用（例如，经济增长是通过提高识字率来促进民主，还是通过改变不同社会集团力量对比来促进民主），而是在总体平均水平上两者之间的数量关系，那么定量方法应该是更好的选择。当然，这种平均水平也是可以建立在一定约束条件之下的，比如研究者仅对 90 年代，或者亚洲地区的数量关系感兴趣，等等。

(2) 因果关系的表现形式。具体地说，就是是否具有多重因果路径、或是否具有 INUS 原因。对于不同的原因造成同样的结果，或同样的原因造成不同的结果这样的情况，使用定量分析方法会遇到较大的困难，因为这意味着一个单一的模型结构无法再能达到目的，或者说，随着样本的增加和减少，定量分析法所得到的结果会出现不稳定和较大的差异。对于 INUS 原因，往往使用定性分析方法更为容易处理。原则上，INUS 因果关系也可以用定量分析法处理。考虑以下 INUS 原因：

$$Y = (X_1 * X_2) \text{ OR } (Z)$$

如果使用回归方法分析，那么正确的模型设定形式应该是

$$Y = \alpha + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot (X_1 * X_2) + \beta_4 \cdot Z \quad (3)$$

其中  $X_1 * X_2$  是  $X_1$  和  $X_2$  的交叉项。

但是，如果 INUS 原因呈现如下形式：

$$Y = (X_1 * X_2 * X_3) \text{ OR } (Z_1 * Z_2),$$

那么正确的回归模型设定形式就不仅应该包括所有在 INUS 中出现的单独项 ( $X_i$  和  $Z_i$ )，还应该包括  $X_1$ - $X_3$  之间的所有两两交叉项，以及  $X_1 * X_2 * X_3$ ，以及  $Z_1 * Z_2$  交叉项。这样一来，不仅大大降低了估计的自由度，对样本量的增加提出了极大的要求，而且在实际中，这种包含多项交叉项模型的估计结果往往并不令人满意。在这种情况下，定性分析方法不失为一个明智的选择。

### (3) 必要/因果关系

必要/因果关系是一种确定性的关系，因而适用于定性分析方法而不是定量分析方法。不过，近年来一些方法研究的进展也开始研究如何用回归方法判断必要关系，尽管这方面的进展并不是特别明显 (Braumoeller and Goertz 2000)。

## 3. 样本选择

对于定量分析方法而言，样本量越大，解释变量  $X_s$  之间的差异越大越好，因为这样会提高定量分析（例如回归分析）结果的统计性质。并且，定量分析注意解释变量样本选择的差异性和随机性，对被解释变量  $Y$  的具体取值并不予以特别的关注。相反，对于定性分析方法而言，其对  $Y$  的选择十分在意，并且往往只会选取那些  $Y$  呈现具体特定值的样本。

研究者指出，如果研究目的是希望考察必要原因的话，那么针对  $Y$  的具体取值而选择样本不仅不会造成估计偏差，而且还是必要的 (Bennet and Elman 2006; Braumoeller and Goertz 2000)。例如，如果要考察  $X$  是否是  $Y$  的必要条件，就意味着需要考察如下条件是否成立： $P(X=1|Y=1)=1, P(Y=1|X=0)=0$ 。这一条件说明，我们必须考察所有  $Y$  取值为 1 的样本，而对于  $Y$  取值是否为 0 的样本并不用在意。

一个例子是对革命爆发的研究。如果要考察某因素是造成革命爆发的必要条件，那么就on应该收集那些革命爆发的所有样本，而对于那些没有爆发革命的样本则没有必要选取。例如 Skocpol (1979) 针对俄国、法国和中国三国爆发社会革命的分析。虽然她也考察了其它没有爆发革命的国家的案例，但她举这些案例的目的不是为了鉴定必要条件，而是为了和爆发革命国家的案例进行对比和说明。

以上的讨论说明，不存在一个先验的标准规定定量（性）分析方法优于定性（量）分析方法。同样地，也不存在一个先验的标准规定某一定量（性）分析方法优于另一种定量（性）分析方法。究竟选取什么样的方法，取决于各种考虑因素。另外，各种分析方法可以用来检验理论的正确性，或者说，是对理论的证伪。但是，我们不能指望通过这些分析方法本身来发现理论。要得到一个理论，必须依赖于这些方法之外的知识、兴趣和经验。

## 参考文献

- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation". *American Economic Review*, December, 91, 1369-1401.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution", *Quarterly Journal of Economics*, 118, 1231-1294.
- Bennett, Andrew, and Colin Elman. 2006. "Qualitative research: Recent developments in case study methods." *Annual Review of Political Science*, pp. 455–76.
- Braumoeller, Bear F. and Gary Goertz. 2000. "The Methodology of Necessary Conditions." *American Journal of Political Science* 44: 844-58.
- Campbell, Donald and Julian Stanley. "Experimental Design." Houghton Mifflin Company. 1966.
- Dawid, A. P. 2000. "Causal Inference without Counterfactuals", *Journal of the American Statistical Association*, Vol. 95, No. 450: 407 – 448.
- Ertman, Tomas. "The Birth of Leviathan: Building States and Regimes in Medieval and Early Modern Europe". Cambridge: Cambridge University Press. 1997.
- Gujarati, Damodar N. "Basic Econometrics" (4 edition). McGraw-Hill/Irwin. 2002.
- Harrison, G. J. List. "'Field Experiments", *Journal of Economic Literature*. December 2004, Vol. XLII, pp. 1009–1055.
- King, Gary; Keohane, Robert; Verba, Sidney. "Designing Social Inquiry: Scientific Inference in Qualitative Research". Princeton: Princeton University Press. 1994.
- Lieberson, Stanley, "Small N's and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases," in Charles Ragin and Howard Becker (eds.), *What is a Case? Exploring the Foundations of Social Inquiry*. Cambridge: Cambridge University Press, 1992, pp. 105-118.
- Mahoney, James, "Strategy of Causal Assessment in Comparative Historical Analysis", in James Mahoney and Dietrich Rueschemeyer (eds.), *Comparative Historical Analysis in the Social Sciences*, Cambridge: Cambridge University Press. 2003.
- Skocpol, Theda. "States and Social Revolutions: A Comparative Analysis of France, Russia, and China." Cambridge: Cambridge University Press. 1979.