

社会科学中的因果分析*

彭玉生

提要:休谟提出的因果问题影响了两个半世纪的科学哲学讨论,即人类如何从有限经验观察推论必然因果关系?密尔从方法论角度阐释了因果推论的逻辑,指出只有在其他因素相同的条件下,经验相关才能推论因果,为现代随机分组试验设计奠定了基础。实验方法在社会科学中应用有限,更为常用的是观察数据统计分析和个案数据定性分析。复杂统计模型用统计控制模拟实验控制,也能有效检验因果命题。定性分析是理论与经验的反复碰撞,有利于提出新理论或新假设。本文强调,因果解释一定包括机制解释,而机制解释一定是理论解释。因果理论不是对经验相关的简单归纳,而是思想飞跃。检验因果理论的基本逻辑是:从理论演绎因果命题,再演绎相关假设,然后用统计数据检验之。最后讨论整合理论、定性、定量的三点定位原则。

关键词:休谟问题 密尔逻辑 随机试验 统计控制 个案 三点定位

科学研究的终极目标是理论解释,对现象的描述和分类是前科学阶段。所谓科学理论,是关于超脱具体现象的一般性概念和命题系统,不仅能够说明是什么,还能够说明为什么。如牛顿力学,解释苹果垂直落地、行星运行轨迹等。社会科学虽然有很大的特殊性,描述性研究占据很大份额,但理论解释仍然是其最高目标(Merton,1968)。

社会科学的理论解释包括两类风格:因果解释(explanation)和意义阐释(interpretation)。解释性(因果解释)研究力图揭示社会现象发生的原因,而意释性(意义阐释)研究则阐释行动的文化意义。我认为,因果解释和意义阐释并不矛盾,相反,意义互动恰恰是社会因果机制的特点。所以,无论是解释性理论,还是意释性理论,因果命题都是必不可少的。除了少数历史社会学家声称只做历史描述、揭示个案的特殊性、避免因果分析(如Bendix,1956),纯粹的描述性研究并不多见,众多学者致力于对现象的因果解释,甚至寻求普遍性因果定律(如

* 作者曾在上海大学社会学系、中国人民大学农村发展研究院和山东大学哲学与社会发展学院的研讨课上演讲本文,感谢与会的学者、学生。黄宗智、刘杰、刘玉照、周立、周怡、田晓丽、阮丹青和许建明教授在文章写作过程中提出了宝贵意见,在此致谢。

Moore,1966;Skocpol,1979;Kiser & Hechter,1991),或者普遍性社会机制(Elster,1999;Hedstrom 2005;Gross 2009)。

社会科学的因果分析比自然科学的因果分析更为微妙复杂,源于两方面原因:其一是社会科学研究对象本身的特点所致。人的自我意识及行为任意性,导致社会现象具有更大随机性,这使社会现象的因果链条变得更加扑朔迷离,因果推论更为困难。例如,两个台球相撞而发生动能传递是一纯粹力学现象,可用力学原理精确描述。两个人发生冲突,可以表现为肢体冲突,权力的较量,也可以表现为纯符号冲突,所谓君子动口不动手;而且肢体冲突和权力较量也是符号互动,因为都涉及对社会情景的界定和文化规则的解释,这就使社会现象的因果链条极为复杂多变,不确定因素增加许多。其二是社会科学研究方法上的局限性。实验方法是因果推论的惟一可靠方法,这是目前科学界的共识,但实验方法在社会科学中应用极为有限。常规而言,社会科学家只能在自然状态下观察研究对象和收集资料,通过定量统计分析或定性个案分析来探索因果关系。定量方法以统计控制模拟试验控制,区分规律性和偶然性因果关系,估算净相关,从而检验因果假设。定性方法基本上是理论和经验资料的反复碰撞,一般不能用来检验因果假设,但以分析事件过程、阐释文化意义、提出新命题为优势。两类方法各有所长,二者的结合则能取长补短。

本文焦点是社会科学因果分析的方法论,旨在讨论各种因果分析方法共同逻辑和社会科学应用中的误区,澄清理论、因果命题和相关分析的联系。下文首先回顾因果问题的哲学讨论,方法论逻辑,探讨统计模型与个案分析的优劣,强调理论抽象在因果解释中的作用;最后讨论理论、定性、定量相结合的三点定位原则。

一、休谟定义、密尔逻辑与随机组设计

(一) 休谟定义

学术界讨论到因果关系,不免要提到18世纪哲学家大卫·休谟提出的因果问题或归纳问题(Hume 2005/1739,1999/1748),因为这是人类认识论的软肋。休谟的疑问左右了其后200多年的哲学和科学哲学讨论。我们先看他对因果概念的剖析。休谟认为,所谓因果,有三层含

义。第一是原因和结果事件之间在时空上毗连(contiguity),时空联接是因果关系的先决条件。如果两个时空相距很远的物体产生了因果作用,那么其间必然存在某种因果链条的衔接。第二是时间顺序(succession),即因先果后。第三是必然联系(necessary connection),即因果现象相伴而生,有其因必有其果。休谟指出,这三要素中,前两条是必要因素,而第三条是最重要的(Hume, 2005/1739:61-65)。这一条也是最具争议的。

关于第三要素,休谟本人用了两个术语,一是必然联系,二是恒定关联(constant conjunction)。二者在语义上有细微不同,恒定关联强调经验伴随现象的重复性,用现代统计学术语就是统计相关。而必然联系则是超越感官经验的,是人类对恒定关联的心理印象。这种定义上的微妙矛盾蕴含了经验主义的基本问题(Beauchamp & Rosenberg, 1981)。一方面,休谟强调,因果联系不同于逻辑推论,一定来自经验;我们不能通过分析原因事件的特性而推导出它的结果(Hume, 1999/1748:109-110)。另一方面,他指出,我们的内外感官都不能直接观察到事物之间的联系:“总的来说,世界上没有任何自然的联系(connection)是我们能够感知的。所有事件都是离散的。一事件跟随另一事件之后而发生,但我们永远无法观察到它们之间的联系。他们似乎相伴而生(conjoined),但并无联系”(Hume, 1999/1748:144)。

休谟认为,因果概念只有一个来源:我们对经验关联现象的重复观察。当我们一次又一次观察到两类现象总是相伴而生,“我们便认定二者之间有某种联系,某种力量,使其中一现象能够屡试不爽地、确定地、必然地产生另一现象”(Hume, 1999/1748:144)。休谟用台球相碰为例来说明:一只运动的台球撞击另一只静止的台球,后者由静变动。第一次看到这个试验,我们只看到前一只球的撞击和后一只球获得速度,至于动能如何传递,我们永远看不到。多次重复这个试验之后,我们觉着二者有某种必然联系,并能预知台球撞击的结果。所有关于必然关系、作用力、能量的概念都来自对恒定关联的心理印象,没有任何外感内省的根源(Hume, 1999/1748:145)。休谟所谓“作用力”应泛解为因果作用机制。

所以,因果判断,以至我们关于自然界的一切知识,都是基于人类对重复性关联现象的归纳。然而,休谟认为,无论对两个事物前后相随重复观察多少次,我们也无法肯定明天它们还会前后相随(Hume,

1999/1748:113-118)。如何从有限的经验观察归纳出必然性的因果判断?这就是休谟的“因果问题”或“归纳问题”。这个认识论问题启发了许多大哲学家,包括康德、罗素、波普尔,引起一系列哲学讨论并延续至今。

休谟处在启蒙运动晚期,也是牛顿力学鼎盛辉煌的时代。其时,休谟的同时代人,受牛顿力学的鼓舞,认为牛顿发现了颠扑不破的绝对真理,对人类认识能力也深信不疑。然而,牛顿强调,万有引力是我们看不见的力量。受牛顿影响,休谟以其冷静的逻辑思考和天才洞识,强调因果关系如同力和能量的概念一样,是我们永远看不见的。就是说,因果机制是永远打不开的黑匣子:“宇宙间万象变换不止,物物相随无缝无隙;但推动这架机器的能量和力却深藏不露,是人类感官永远不能发现的”(Hume, 1999/1748:136)。

到此,我们看到休谟因果概念的张力。一方面,休谟将因果概念剖析得干干净净,只剩下经验关联。但是,经验关联不具备必然性,必然性是人类对经验关联的心理印象,是人类的思维习惯。另一方面,休谟强调必然性是因果概念的最重要因素,是因果推论的基础,没有必然性,就无法从原因推诸结果(Hume, 2005/1739:129)。用今天的话说,休谟所谓必然性联系就是普遍性因果规律和一般性理论。规律必须具备普遍性,至少一定程度的普遍性,否则就不能称其为规律。恒定关联与必然联系之间的张力就是经验与理论之间的矛盾。如果人类放弃对普遍性定律的执著,休谟的归纳问题也就不存在了。但是,人类不满足于有限理论,而是不断地探索更具普遍性的统一理论。

解决因果概念的张力有两种方案:一是将其下放为纯粹经验概念,二是将其上升为纯粹理论概念。下放为纯经验概念就是说,所谓因果,不过是经验关联的代名词。比如大哲学家罗素(Russell, 1903)认为,休谟本人的经验主义立场不够彻底,若将经验主义推至其逻辑终点则必然得出结论:既然因果概念既非经验、亦非逻辑,干脆将其从科学术语中清除出去。现代统计学奠基人皮尔森(Pearson, 1900)更缘此提出用相关分析代替因果概念。

时至今日,因果概念不仅没有被现代科学摒弃,其地位反而更牢固。我认为,因果概念与科学理论密不可分,因而绝非毫无意义。摒弃因果概念,则有摒弃理论之嫌疑。科学知识不仅来源于经验感官,还需要理论抽象。对因果机制解释,追根求源来自理论或常识(即理论和

经验知识的沉淀)。在日常生活中,因果概念就更必不可少。医生诊病救人,法官定罪量刑,都离不开因果分析。

所以,我更趋向第二种方案:将因果概念上升为理论概念。第二种方案的代表人是康德。受休谟归纳问题的震撼,康德(Kant, 1998/1781)提出哲学二元论,即人类知识有两个来源:感觉经验与纯理念;纯理念是超验的,与经验世界之间有一道不可逾越的鸿沟。康德的纯理念不仅包括逻辑概念,也包括因果概念和物理规律。他似乎认为,人类将先验的因果规律(如万有引力)强加给了经验世界。

本文认为,因果是理论概念。一方面,我认同休谟的观点,人类的因果知识都是基于对经验相关的归纳。另一方面,人类因果判断不是简单的归纳而是一种抽象。人类思维不满足于“有A必有B”的判断,还要解释为什么,解释因果机制,而机制解释是理论抽象。所以,因果概念不应下放为纯粹经验相关,而应上升为理论概念。但是,因果概念也不是康德的纯理念般与经验世界毫无关联。感觉是主观的(色彩、声音、冷热、时空等),但我们对外界事物的变化和关联的印象却是相对客观的,这是检验因果理论的认识论基础(关于因果理论的产生与检验,容后文续论)。

从方法论角度来看,休谟并没有阐述因果分析的具体操作方法和步骤。关于判定因果关系的方法论,他在时空毗连、时间顺序和恒定关联的基础上又增加了五条(Hume 2005/1739:135-136)。其中第四条是,有其因必有其果,有其果必有其因。换言之,如果A与B有因果联系,那么有A则有B,并且无A则无B。第五条是,假如不同原因能导致同样结果,那么这些诸多原因必有某种共性,且是真正原因。与上文讨论的从经验推论必然的认识论问题不同,这两条原则是方法论问题,但表述显然太狭窄,对现代科学方法没有产生任何影响。第五条基本上是单因决定论,可能受牛顿万有引力的影响,也可以解释为对统一理论的追求。

第四条有现代实验方法的雏形,但其表述有两个局限性:其一,过于狭窄,可能与当时统计学欠发达有关。用现代统计学表述,两个事件的关联定义为:有A则B发生概率较高,无A则B发生概率较低。换言之, $P(B|A) > P(B|A')$ 。休谟的第四原则(有A则有B,无A则无B)只是特例,即: $P(B|A) = 1$,且 $P(B|A') = 0$ 。其二,从经验相关推论因果是有条件的:只有在其他因素相同条件下,我们才能从经验关联

推论因果关系。休谟本人没有说明这些条件。时隔一个世纪,约翰·斯图亚特·密尔(John Stuart Mill)深入讨论了这些条件,严格定义了因果归纳的逻辑,为现代实验方法打下了根基。

(二) 密尔逻辑

像19世纪许多哲学家一样,密尔(Mill 2002/1848)强调归纳乃是一切经验知识的基础。他从方法论的角度总结了因果归纳的基本逻辑,被后人称作密尔逻辑。密尔逻辑并不从认识论角度解决休谟的归纳问题,而是讨论因果推论的方法论原则,我认为其重要性在于指出了因果归纳的限制性条件。

密尔提出了四种方法:求同法、求异法、剩余法,以及共变法。其中,前两种方法应用广泛,剩余法用途有限,而共变法是求同法和求异法的结合。

简单说来,求同法是:如果两个或多个个案除了A和B之外没有任何共同之处,那么A和B必有因果联系。求异法是:如果两个个案各方面都相同,但一个个案有A有B,而另一个个案无A无B,那么A和B必有因果联系。下面是密尔本人对求同法和求异法的解释。

第一原理(求同法):“如果所研究现象的两个或多个案例只有一个共同特征,那么这个共同特征就是所研究现象的原因(或结果)”。密尔解释道,这个“发现、证明自然规律的方法来自以下公理:任何可以被排除而不影响所研究现象的特征,即当其消失后而现象依然存在的特征,与该现象一定无因果联系”(Mill 2002/1848:255)。当所有特征都被排除后,不能被排除的特征一定是原因。

第二原理(求异法):“如果所研究现象发生的案例和所研究现象未发生的案例各个方面都相同,但只有一个特征不同,这个特征被前者所具备却不被后者所具备,那么,这两个案例惟一不同的特征就是所研究之现象的结果或者原因,或者原因的不可缺少的一部分”。求异法来自以下不证自明的公理,即“若一前件一旦消失就招致所研究现象消失,那么该前件必是研究现象的原因;如果其他前件都不变而只有一个前件消失,那么任何随其消失的后件,必定是那一个前件的结果”(Mill 2002/1848:256)。

与休谟明快的语言风格相比,密尔的语言晦涩难懂,翻译成中文更加拗口。斯高切波和索莫斯(Skocpol & Somers, 1980)对密尔逻辑做了

图解,一目了然,我基本照搬,见图1。求同法可以总结为:在总体不同条件下,关键一致是因果。求异法可总结为:在总体相同的条件下,关键区别即因果。

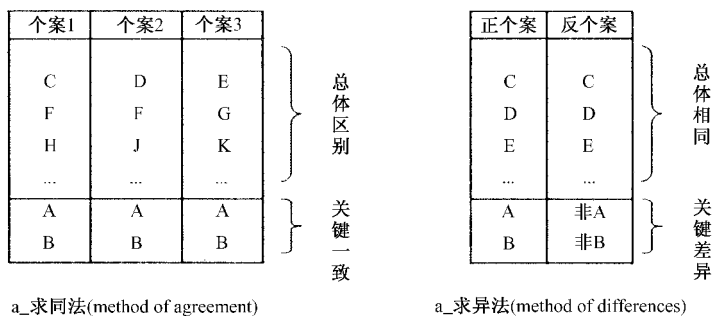


图1 密尔逻辑图解

密尔本人称这些方法为实验方法,因为总体相同和总体区别只有实验室能够达到,现实中只能近似。自然现象和社会现象往往是多因的,并且这些原因又是彼此交织关联的,那么要想确定A是B的原因就必须将它们从其他因素中剥离出来,将其他因素予以“控制”。实验逻辑是分析逻辑,即肢解、分离现实中具体的、复杂的因果过程,抽取简单的、一般性的因果联系。求同法的“总体区别”条件不可能严格操作,是日常生活中常用的直观的、辅助性的推理工具。密尔自称“最完善”的求异法更受科学家青睐,其“总体相同”条件就是现代经济学的口头语“其他因素不变”,或者社会学的“可比性”。

我认为,密尔对归纳逻辑的定义是严格无误的,虽然不解决休谟的归纳问题,但绕开了经验关联与必然关系之间的矛盾。既然从个案推论一般是有条件的,那么因果的“必然性”也就有了限制条件。现实经验世界中鲜有必然的因果现象。从本体论角度理解密尔逻辑,这是因为因果关系是有条件的,即:其他因素不变。对于现实中多种因素交织综合产生的因果现象,因果陈述还可以是必然性的,但必须附带条件。比如谚语道:“瑞雪兆丰年”,其条件是没有霜冻,没有冰雹,没有蝗虫,等等。有时条件可以多到因果陈述脱离现实意义,比如,勤奋读书一定能考上好大学,假如你智商在90%以上,身心健康,没有不良习惯,不交损友,老师好,学校好,不是农民女儿,家境好,考试当天不发烧,不落

水沟……等等。

现实生活中看似单因的、具有必然性或接近必然性的现象也是有条件的,只不过这些条件蕴含在我们对日常现实的定义中,不言而喻。比如苹果熟透了必然落地,假定没有龙卷风,没有人摘食,苹果梗没有病变等。太阳明天早上要升起,假定没有日食,没有彗星撞地球,没有黑洞吞噬太阳系,万有引力不消失等。从本体论角度看,如果因果规律是有条件的,绝对的因果规律也就不绝对了。人类可以追求完美的、放之四海而皆准的理论,但是一定要明白,现实的因果规律都是有条件的、相对的。

密尔逻辑在日常生活中运用十分广泛,只是不严格而已。比如,约朋友吃午饭后都拉肚子,我们判断是餐馆有问题,这是求同法。看了两家医生后效果不同,我们认定大夫医术有别,这是求异法。执法人员和侦探通过比较案件之间的共同点和不同点来寻找线索,但未必能呈堂作证。当艾滋病在美国出现时,医生发现患者都是同性恋,所以怀疑这个怪病与同性恋有关,这是求同法;后来发现了 HIV 病毒,才找到真正元凶,同性恋只是偶然因素。

(三) 费舍实验设计

虽然密尔逻辑的表述是个案比较,但个案比较只在某些物理和化学实验中才能满足“总体相同”。因为这类现象的相关因素较少,因果关系相对简单,容易控制。例如,在空气中,鹅毛与铅球不可能同时坠地,因为除了重力之外,还有空气阻力。只要排除了空气阻力(真空管),鹅毛和铅球一定同时坠地。月球和木星的引力可以忽略不计。再如,用同一个铅球从不同高度坠落,试验速度与能量(对地面打击力)之间的关系,空气阻力都可以忽略不计。牛顿力学之所以能够率先发展,其中一个原因可能就是力学现象的因果关系简单,有些规律通过自然观察(比如行星轨道)就能认识。

但大量生物现象和大多数社会现象不仅是多因的,并且带有随机性。对带有随机性的现象,即使在实验室条件下,个案比较常常也难满足总体相同。所谓随机性就是说,能影响某事件的因素非常之多,有些是我们知道的,有些是我们不知道的,并且还有许多细小的偶然因素,根本不可能也无法控制。所以,对许多生物和社会现象来说,自然观察很难发现规律性,即使在实验室,个案比较也难按密尔逻辑操作。比

如,在实验室制造两个一模一样的小白鼠亦非易事,除了先天因素,还有后天因素,虽是同胎所生,对药物的反应也不一定完全一样。与无生命的物理运动相比(如台球相撞),动物的行为带有更大任意性,但大体上受基因控制,遵从弱肉强食的丛林法则。人的行为就更复杂了,除了先天基因密码,还有社会文化密码,经过大脑对情景做出判断、解释,再以其有限理性做出决策。并且,作为有高等智慧的人类,还可以决定逾越游戏规则,极端到作奸犯科、造谣撒谎。所以,人的行为虽然不是完全随机,但偶然因素大大增加。

如何控制偶然因素的作用,有赖于近代统计学的发展。统计学家费舍(Fisher, 1935)完善了随机化实验设计。这里的随机化,就是将试验对象随机分配到控制组和实验组。根据大数定理,个体的各种不可控差异,在随机分配过程中被平均了,即平均而言,两组具有“总体相同”。例如,制药厂测试新药,找一群符合条件的实验对象,然后随机分为两组,给一组施用新药(实验组),另一组施用安慰剂(控制组),最后比较结果,如癌变率、胆固醇水平等等。因为两组平均而言没有差异,所以所有结果有任何差异,都归因于新药。稍微复杂一点的试验设计,可以先将实验对象以根据某些变量(如性别)分成区(blocks),然后在各区内再随机分控制组和实验组,目的是进一步减少控制组与实验组之间的区别。

随机组设计使密尔逻辑具备了广泛的可操作性。统计学对因果关系表述为:在相等条件下,如果A发生,则B发生的概率提高,或者X变化导致Y平均值的变化。因果关系的必然性不表述为个体事件,而表述为群体概率或平均值和随机组试验的可重复性。比如,我们说抽烟有害健康,并不是说每个烟鬼都患肺癌,而是指平均而言,抽烟提高癌变率。酒后驾车易致交通事故,但并非每次酒后驾车必肇车祸。对于个体而言,必然性只是偶然性趋于零的特例,比如赤手停留在100摄氏度滚水里必然烫伤,无一例外。严格说来,随机组实验是因果推论的“金本位”,是目前科学界公认的可靠方法(Fisher, 1935; Holland, 1986; Sobel, 1995, 1996; Winship & Sobel, 2004)。广义地说,因果推论指根据实验结果推论某种普遍性因果关系存在。至于这种因果关系的机制是什么则超出了经验归纳的范围,属于理论解释。

二、统计模型与个案分析

在社会科学中应用实验设计虽然不是绝对不可能,但应用有限。比如,我们没办法给一组实验对象注射 200 毫升“社会资本”,然后看有什么反应。通常研究者都是通过抽样调查在自然的环境中进行观察和搜集资料,然后用多元统计模型来研究经验资料,并通过统计控制来模拟实验室控制。并且,宏观历史研究的对象,一定是个案,不可能运用随机设计。所以,社会科学的因果分析常用定量方法(统计模型)和定性方法(比较个案研究)。

(一) 统计模型和统计控制

统计分析的基本逻辑是将数据分割为系统模型和随机残项。所谓模型就是理论解释部分,而不可解释部分则被归入随机残项,即:

数据 = 模型 + 随机残项

如广义线性模型(GLM)定义为: $Y = \hat{Y} + \varepsilon$

其中 $\hat{Y} = g^{-1}(X\beta)$ LM 特例为 $\hat{Y} = X\beta$

用社会科学通用语言,统计模型可以解释为:

现实 = 理论 + 误差

就是说,经验现实由理论和误差两部分构成。现代社会科学,包括宏观历史研究,基本上放弃了决定论,即便是最完美的理论也不能排除偶然因素造成的误差,并且人类认识能力有限,理论本身还有谬误。

模型(理论)部分由核心自变量和控制自变量组成。统计分析中的“其他因素不变”包括控制变量和随机残项两部分。理论上说,控制变量包括所有统计显著的自变量,剩余所有不显著的因素都被纳入残项,并假设残项完全随机,即与模型中自变量无关。假设残项完全随机,类似实验设计的随机分组。类似程度取决于对相关变量的控制是否充分。实践中,我们只能控制那些已知的相关因素;凡是未知的,或者已知而无数据的因素,都被假设与核心自变量无关而纳入残项。假设成立与否,有赖于研究者对相关因素的认识。所以,从绝对意义上说,只有随机化实验数据才满足“总体相同”的条件,观察数据的统计控制只是近似(Blalock, 1980)。

社会科学不关注的随机残项,在幕后默默影响研究人员的因果推

论。社会科学理论解释力低,就是因为随机因素作用大。解释个人社会行为的线性模型, R^2 一般不超过 0.5,就是说,未被解释的随机因素作用占大半。相比之下,生物学的 R^2 一般都在 0.8 或 0.9 以上,而物理学则能精确到误差可以忽略不计。但这并不影响社科研究人员对自己理论模型的自信,坚信将未知因素归入随机残项,比古人将不理解事件归结为鬼神意志或祖坟风水更有道理。易经八卦理论历时几千年而不衰,今天又以国学面貌走入大学课堂,在笔者看来,是因为其占卜功用事实上模拟人类事物的偶然性,是我们祖先对付不确定性的手段。

关于控制变量的多少有两个互相矛盾的指导性原则:一方面我们应该尽可能多地控制应该控制的变量,另一方面还要避免控制不应该控制的变量。应该控制的变量有两类,第一类是那些对因变量有重要影响的因素,无论与你关心的核心自变量有关与否,都应该包括在模型中,因为只有这样才能建立一个可靠的模型。例如,如果研究收入在性别上的差异,即使年龄与性别没有关系,研究者也应该控制年龄和年龄平方(代表工作经验)。

第二类是那些既影响因变量又影响核心自变量的因素,控制它们才能消除伪效应(spurious effects)。例如,笔者(Peng 2004)在分析宗族网络与企业发展之间关系时,控制了县域,因为宗族网络和乡镇企业都集中于东南沿海地区。如果不控制县域,宗族与企业数目之间的相关可能只反映它们空间分布的耦合。控制了县域,宗族与企业数目之间的系数就只反映每个县域之内村与村之间的比较。

经济学界有一项研究比较有意思,就是孪生样本的收入研究,通过巧妙设计样本,控制无法测量的先天因素。一般抽样数据对教育回报的估值可能偏高,因为先天因素被归入残项。我们知道,先天因素肯定影响教育,并且可能直接影响收入。但先天因素无法测量,智商和情商都是先天和后天两种因素的结果。经济学家从 20 世纪 70 年代就开始收集、分析同卵双胞胎数据,直到今天兴趣依然不减。比较孪生兄弟/姊妹教育水平的区别对收入的影响,就可以排除先天因素影响。

一般收入模型界定如下:

$$\ln Y_{ij} = \beta_0 + \beta_1 E_{ij} + \beta_2 A_{ij} + \beta_3 F_{ij} + \gamma X_{ij} + \varepsilon_{ij}$$

其中 E_{ij} 是第 i 对双胞胎($j = 1, 2$)的教育年限, A 指先天因素, F 是家庭因素, X 是其他控制变量如工作经验、婚姻状态等, ε 是随机残项。其实我们无法测量先天能力和家庭因素,但因为孪生兄弟/姊妹的先天

能力和家庭因素是等值的,所以:

$$\ln Y_{i1} - \ln Y_{i2} = \beta_1 (E_{i1} - E_{i2}) + \gamma (X_{i1} - X_{i2}) + (\varepsilon_{i1} - \varepsilon_{i2})$$

这就大致排除了先天能力和家庭环境的影响,并且年龄、性别、种族因素也都被排除了。由于资料来源和技术操作的不同,经济学家对教育回报的估值有相当大浮动,但大体上每多一年教育能“净”增收入8%至10%左右。多数研究的结论是,控制先天因素与否对教育回报的估值影响不大(Miller et al. ,1995; Bonjour et al. 2003)。

这种煞费苦心的巧办法毕竟是凤毛麟角。常规统计分析中只能量力而为,能控制多少是多少。定量研究有一条不成文的行规:只需要控制那些已知的相关因素,次要并与核心变量无关的因素归入随机残项;至于可能重要而未知的因素,留给后人去发现。所以,如果你要批评别人漏掉了某因素Z,那么必须说明为什么Z不仅影响Y,而且影响X,因而导致参数估值偏高。若说不出原因,这种批评被称为“廉价的挑刺”。如果你指责别人的系数“可能”是伪相关,但又说不出遗漏的变量是什么,这种批评就比较蠢了。

第三类可控因素是中间变量,即那些介于因变量和核心自变量之间的变量。如果控制了中间变量,核心变量的参数估值不包括间接作用。以农业产出为例,新技术的应用(例如杂交水稻、温室蔬菜等等)是教育影响收入的中介因素,控制了技术应用,教育作用可能就消失了(或减弱),但这不说明教育无用,而说明其作用是间接的。再比如,控制子女教育,父母教育对子女收入的影响消失,这不说明父母教育没有作用,而说明家庭通过投资子女教育,影响其收入。控制中间变量还可以作为检验因果机制的一种策略,容后文详述。

另一方面,统计实践说明,加入次要的控制变量一般对核心变量的参数估值影响不会太大,并且控制不当还会干扰估值准确性。以下两类变量不宜控制:(1)因变量的不同测量,或者因变量的一部分;(2)因变量的结果。控制因变量的不同测量会抑制其他自变量的作用。比如,研究小额贷款对农民收入的影响,不能将粮食的产出作为控制变量,因为它是农民收入的组成部分。控制因变量的结果是统计分析一大忌讳,因为统计不分辨因与果。例如,研究父代教育对子代教育的影响,不应控制子代收入。这些都是统计学常识,毋庸赘言。

除了变量控制的难点,因果定义的第二要素——因先果后,在统计分析中也变得模糊、复杂。人的行为可以是对尚未发生事件的预期的

反应。比如,生产过程中,我们认为投入品是因,而产出品是果。但是现代生产过程是根据需求进行的,因而产出成了投入的原因。所以,变量互为因果,或者经济学家所谓的内生性问题(endogeneity)又成了定量研究者感到棘手、头痛的问题。于是,一大批统计学家致力于研究时间序列模型、工具变量等等,为提高参数估值的准确度绞尽脑汁。

鉴于统计分析在控制变量(伪相关)和时间顺序(内生性)两方面的难点,统计学家对非实验设计数据的因果推论持有保留态度,因为太多不确定因素取决于研究者的知识和技巧。统计模型做的再精巧,也只是对密尔逻辑的近似模拟,只有在严苛条件下,因果推论才有效(Blalock, 1980; Rosenbaum, 1984; Sobel, 1995; Pearl, 2000, 2009)。但至少,统计学家一般对自己发现的相关系数或回归系数还是有信心的,尽管参数估值未必精确。逻辑上说,统计相关未必直接证明因果,但是有因果一定有净相关。所以,多元统计分析仍不失为一种检验、证伪因果命题的有效方法。

(二) 比较案例分析

密尔自己曾指出,他所总结的逻辑不适合在“道德科学”中应用,因为人有自由意志,社会现象都是多因的等等。他举例说,要证明自由贸易能促进经济发展,必须找两个在自然禀赋、人口素质、法律、文化习俗各方面都相似的国家,一个推行自由贸易政策并且经济繁荣,另一个限制贸易并且经济落后。但在现实中不可能有两个这么相似的国家,即便存在,也不可能在各方面都类似而仅仅在贸易政策上有区别,因为贸易政策是有先决因素的,如果这些先决因素不同,那么总体相似的条件就不成立(Mill, 2002/1848:575)。

尽管如此,密尔方法依然被广泛地应用于社会科学研究,特别是宏观历史比较研究。比如,斯高切波(Skocpol, 1979)声称她要严格地应用密尔逻辑分析比较法国革命、俄国革命和中国革命。显然,这三个案例不可能满足密尔逻辑的条件。如果说统计控制是对实验控制的近似模拟,那么个案分析的“控制”最多是对统计控制的遥远模仿,最多只关注“理论上相关的”因素来确定“总体相似”或“总体区别”。有时,历史社会学家对某事件做时间前后的比较,因为是同一个案,所以有一定程度的总体一致。

偶尔,质性研究者能幸运地发现非常相似的案例,模拟“自然实

验”的效果,以下是两个范例。

范例1:李静君(Lee,1995)比较分析了香港和深圳的两个工厂。这两个厂有同一个老板和同一个管理团队。但两个厂管理工人的策略却大相径庭,深圳厂有“专制”特征,而香港厂则采取比较民主的管理策略。排除了产权、管理、技术及产业等因素之后,李静君将劳工待遇和劳动条件等差异归因于劳动力市场的构成、组织和环境。

范例2:寇古、赞德(Kogut & Zander,2000)比较东、西德两家蔡司厂,声称这两家厂构成“自然实验”,因为有共同历史、同样品牌、同样技术、都是战后重建等等,但一个厂是计划经济,另一个是市场经济。那么,作为计划经济的东德蔡司厂是否缺乏创新能力?通过收集、比较战后两家厂的专利数量,他们发现东德蔡司厂的创新能力并不逊色于西德厂,甚至超过后者。这一经验观察证伪了所有社会主义企业都缺少创新能力的命题。同时,他们发现东德蔡司厂专利更具多样性,而西德厂的专利则比较集中。他们的解释是:东德厂科研听命于中央计划不断变换的战略目标,西德厂受市场竞争压力而集中于优势专项领域。

定性研究如此关注个案可比性的并不多见。即使这两个范例,距密尔逻辑的总体相似还相去甚远。仅制度环境不同一项,就漏掉了许多潜在的相关因素,因为制度的内容太多。其实,这几位作者都是个案导向,没有声称要证明一般性因果理论。李静君无意证明家庭主妇构成的劳动力市场永远比打工妹劳动力市场更加民主平等。寇古、赞德也无意证明市场竞争一定使企业技术研发更加专一,而中央计划一定增加企业技术发明的多样性。

一般而言,对自然状态下的个案分析,密尔逻辑的有效性有限。它不仅不能证明恒定关联(即在相同条件下能够重复的关联),甚至不足以证明个案自身的具体因果联系。实践中对具体因果现象的分析除密尔逻辑外,更需要大量理论知识。例如,一个烟鬼患了肺癌,我们断定抽烟是罪魁,不是因为他有个孪生兄弟烟酒不沾且身体健康,而是因为我们知道尼古丁的危害。一个人酒后驾车撞到树上,我们知道是酒精作怪,不是因为树不长眼。在理论欠缺的情况下,具体现象的因果分析

都有难度。比如,中世纪欧洲黑死病肆虐,人类除了祈祷上帝,束手无策。现代社会出现艾滋病、SARS,科学家能很快找到病源,因为有病毒学理论指导。

《时代周刊》曾报道过一个案例,十分有趣,说明个案资料与理论的关系。一北京记者报道了华北某村干部违法滥权的故事,不料这村干部来头很大,以诽谤罪将记者告上当地法庭,并要求记者赔偿经济损失,论据是自从记者文章发表之后,村里“树不结果、鸡不下蛋”(Time Asia 2000)。官司以记者败诉收场。

在这里,我们不去评议法院判决的缘由,而是分析诉方的证据。从密尔的求异法角度来看,起诉方的论据是历史比较法,即比较同一个案在某事件发生前后的差异,这是历史学家常用的方法。诉方的论据也有“总体相似”的条件,因为记者出现前后,面对的是同一个村庄。“关键区别”的条件也符合:记者出现以前,树结果、鸡下蛋;记者出现以后,树不结果、鸡不下蛋。这个村的历史上大概没有发生过这样的倒霉事。所以,记者的出现与树不结果、鸡不下蛋有脱不掉的干系。北京来的辩方律师大可私下窃笑:这太荒唐,没有一点科学根据。然而,诉方有根有据:因为记者的恶言恶语坏了村里的风水,鬼怒神怨,所以树不结果、鸡不下蛋。这就是因果机制,只是北京来的律师不信祖宗留下的这种“精神”财富罢了。相信科学的北京律师也许指出天气异常、水源污染等因素是导致“树不结果、鸡不下蛋”的原因,并用一套科学理论来解释因果机制。但是,单单从经验证据上看,记者出现和水污染都不能被排除。说到底,诉方论据是否能站住脚取决于你(包括法官)相信哪个理论:风水理论还是科学理论?

所以,与其说个案研究能验证理论,不如说是理论在解释个案。并且,由于个案的代表性无从判断,用个案检验一般命题十分不可靠。具体因果过程往往是多种因素包括偶然因素的交织作用,因此,个案都具有特殊性。即使我们对个案的经验描述和理论解释都准确无误,也无法确定从个案分析中梳理出的因果关系是否带有普遍意义。换言之,我们没有大样本来判断个案所反映的因果关系是否显著。一个没有代表性的、甚至是错误的理论碰巧找到与其一致的随机抽样数据相当困难,但总能找到与其一致的个案数据。所以,除了证伪“天下乌鸦一般黑”之类的绝对命题外,个案分析不能检验一般假设。更何况,社会科学多用“一般命题”类概率论语言,“绝对命题”的表述很少见。

由于比较个案研究很难证明经验相关,所以更重视因果机制分析,即从理论上来解释现象之间的关联。有些人认为定性研究理论程度高,而定量研究理论成分低,这种看法有一定依据。定量研究依靠数据证明核心变量之间的相关关系,对理论的关注点是假设验证,因而用在理论讨论的篇幅相对较少,即使没有理论支撑,也能证明经验相关。而定性研究则要靠理论来论证经验关联,需要更多理论支撑,而没有理论支撑的个案研究更可能是空洞无物的糟粕。本文无意妖魔化定性研究,而是要强调理论思维在定性研究中的重要性。人类的知识、科学的发展都是积累性的,运用现有知识无可厚非。既然理论在个案资料的筛选和分析中是必不可少的,那么个案研究从根本上说就是理论和资料的反复碰撞,其结论即为事后解释。

正因如此,许多个案研究者干脆回避实证主义语言,把研究目标锁定在对具体个案的描述或解释上,而不是证明普遍性因果理论。如华尔街次贷危机的根源是什么?“文革”何以发生?研究者援引或创新理论,闭口不提假设检验。这类研究的结论既依赖经验资料的可靠性,也依赖被援引理论的有效性。如果理论是已被验证和接受的,那么结论就有说服力;如果理论是尚有争议的或者新创的,那么解释只是新的假设而已。

(三) 叙事法与事件过程分析

近20多年来,越来越多的个案研究者,放弃用密尔逻辑来做因果推论的企图,转向事件过程分析。在“从因果到事件”一文中,亚博特(Abbott,1995)称其为一场“静悄悄的革命”。叙事法(narrative)或者详叙法成了个案研究者的前卫术语和口头禅(Denzin & Lincoln, 2000)。所谓详叙,指详细地叙述分析,关注点一般是事件的时间序列(sequence of events)或事件发生链,所以又称“过程事件分析”(孙立平,2000)。叙事法未必排斥因果分析,对具体事件的分析可以是因果解释,也可以是文化阐释。所以,详叙不是单纯的经验描述,还包含理论分析。

有趣的是,叙事法常用一个叫“反事实分析”的分析技巧,是指叙述个案事件发生序列后,假设如果大环境不变,某该发生的事件没有发生,或者不该发生的事件发生了,那么其后的事件序列会有何不同(Griffin,1993)。这显然是对密尔逻辑的模拟,先在大脑里假设总体相

似,再靠理论或常识推理,虚构一个反案例,所以,又被称为思想实验。例如:19世纪末智利曾发生内战,以民族资产阶级失败告终,其失败纯属历史的偶然,其后智利走上依附发展的道路。假如智利的资产阶级革命取得成功,那么智利可能早已跻身发达资本主义的行列。这是一个真实的历史可能性(Zeitlin,1988)。

叙事法对时间序列的关注似乎是对休谟因果概念中的时空毗连的回归。休谟定义毗连为事物在时空上的联接,是因果关系的必要条件,他说:“所有因果都是毗连的,任何间断都会使因果作用无法操作。遥远的物体也会产生相互作用,但仔细观察都会发现它们有一条因果链条,而这些链条相互毗连”(Hume 2005/1739:63)。

事件发生链与因果机制的概念有某种类似,因为机制也是一个过程概念。那么,分析事件发生链是否能够揭示因果机制?我认为,事件发生链与因果机制并非同一个概念。因果机制是具有普遍意义的理论概念,而事件发生链则指具体经验现象的发生过程。仔细观察原因与结果之间的事件序列,只是将经验现象在时空上分解得更细,其所揭示的事件发生链仍然是经验上前后关联,所以并不能揭示因果机制或证明因果关系。如果仅仅靠观察事件发生链就能发现因果机制,科学就太容易了。相反,详细叙事本身有赖我们对因果机制的理解,对相关事件的筛选,对每一条因果链的描述分析,这些都包含理论分析。虽然事件发生链不是因果机制本身,对事件发生链的细致观察应该能够拉近研究者与因果机制的距离。人类知识的积累和科学的高度发展意味着大量现象的因果机制可以通过深入观察和理论分析来解释。

举一简单例子:对苹果落地的过程进行细节的描述(如快速定时拍照)能揭示垂直运动和恒定加速度,但并不揭示万有引力。相反,根据力学定理,科学家可以推论垂直运动和恒定加速度意味着某种恒定力量的作用,又根据万有引力理论来解释恒定力量。对简单的自然现象的描述尚且如此,对复杂的社会现象的描述和文化解释就需要更多的理论和常识(或者文化执见)。社会机制归根结底是必须落实到个人的文化心理过程的,难以直接观察。这不是说社会机制不可知,而是要强调:所谓机制解释,是理论解释;机制证明,是理论证明。^①

① 笔者对质性研究不是内行,对其技术细节不予深入讨论。对此类方法有兴趣的读者请参考专业著作(如 Abbott 2001;Denzin & Lincoln 2000;Abell,1993;Ragin,1987)。

本文在此总结质性研究的三大优点。

优点一,质性分析的关注点是个案本身。统计模型是以发现显著相关关系为目的,其关注点是统计显著的概念和变量,受随机因素左右的个案只是数据点而已,所以这也被称为“变量导向”。变量导向的统计分析不可能详尽地描述事件过程,因为它所能分析的变量数目有限。叙事方法就是因为不满于变量导向的统计方法而提出来的(Abbott, 2001; Ragin, 1987)。个案分析,顾名思义,是个案导向,关注个案本身的故事,无论这故事是特殊的还是带有一般性的。质性研究者的特长是以敏锐的观察力,将个案的事件过程详细叙述出来。

对个案导向的研究者来说,那些被统计模型扔到随机误差项的偶然因素,很可能是重要的而且有趣的故事。想当年李自成打进紫禁城,坐进了金銮宝殿,吴三桂见大势已去,率部归降闯王,路上巧遇逃难家仆,听闻爱妾陈圆圆被刘宗敏所霸,冲冠一怒为红颜,反引清军入关,打败了李自成。这期间,刘宗敏霸占陈圆圆完全是偶然事件,吴三桂路遇家奴也是偶然事件,这些偶然事件却改写了中国历史。退一步说,假设刘宗敏没有霸占陈圆圆,或者吴三桂忍下了这口恶气,那么中国的末代皇帝就不是爱新觉罗的后裔,而是李家后代了。

这一对中国历史产生重大影响的偶然事件并不具备普遍意义,无法概括。统计模型则会忽略这类偶然事件,即使包括在模型中也不会有显著性。但离开了偶发事件,鲜活的历史可能变得索然无味。试想,如果人类历史用几个概念就能解释清楚的话——不管是马克思的生产力和生产关系,还是帕森斯的格局变量,那么人类历史会比行星的历史还单调。而质性研究则可以通过叙事法将鲜活的历史再现出来。

优点二,文化阐释。人类的互动都是符号性的,所以对人类行为的详细描述还需要加入符号意义的阐释,即所谓的深厚描述。文化意义的界定和符号互动不过是社会因果机制中的环节而已,与因果分析并不冲突。但文化意义难以直接测量,所以实地调查和民俗学方法是必不可少的(Geertz, 1973)。定量研究也有各种态度和价值量表,但这些量表必须建立在深刻了解地方文化的基础之上。

优点三,虽然不利于检验假设,质性研究却有助于理论和概念创新。如果现有理论或概念不能解释新经验现实,那么必须创造新概念或新理论。人类社会的发展日新月异,不停地产生新的现象,提出新问题,超越现有理论。比如中国最近30年的飞速发展就挑战了许多西方

现有理论(黄宗智 2007)。深入的实地调查、细致的事件过程分析,是创新理论的必要条件。定量研究必须以既有概念和理论框架为前提,因为所谓变量,不过是对现有概念的测量。对社会现象的描述分类是概念测量的基础,需要深入的民俗学观察研究和细致的个案比较分析。定量分析的创新一般限于现有概念之间的新联系;而实地观察和个案分析则不必囿于现有的概念或理论框架,可以从实际出发,提出真正关键问题,为理论创新提供了更大空间。个案观察也需要理论思维能力,许多个案研究者往往囿于经验细节,见木不见林。如果个案研究者有理论想象力,对事件过程做深入观察、细致描述,并能做出理论解释,甚至提出新的理论假设,那么就能够创新理论。

三、因果机制与科学理论

密尔逻辑定义的因果关系也是纯经验性的,可以理解为不能被其他变量解释的、单向的相关关系。无论实验方法还是统计方法,只能证明经验关联并推论因果关系的存在,并不能说明这种关系是什么。完美设计的随机组实验,也只能证明 A 导致 B,至于 A 如何导致 B 却像一只黑匣子。这就是机制问题。因果概念如果有客观意义(causal realism),一定要包括对因果机制的解释,否则因果只是密尔逻辑所定义的相关概念的代名词。

机制解释和理论思维是人类独有的能力。相关思维是对经验现象的概括,动物界也有类似能力。巴甫洛夫的条件反射试验说明,狗也有相关思维能力,能把铃声与食物联系起来,但它们大概不会去琢磨铃声为什么带来食物,即相关背后的因果机制,只关心摇铃吃饭、吃饱睡觉。因此动物也不会焦虑为什么环境变了之后,铃声不再带来食物。人类思维复杂得多,也不满足于眼前吃饱睡觉,所以创造出许多抽象理论和概念来解释经验相关,如神的意志、善恶报应、阴阳风水,以及万有引力、遗传基因、社会资本等等。在科学高度发展的今天,人类生活越来越多地依赖科学理论来解释经验相关,许多抽象的理论和概念逐步演化成了日常生活的常识,因果思维与经验世界的鸿沟也越来越模糊。而神学和玄学理论仍有市场,因为仍然有大量科学尚未解释的现象或根本不可能解释的纯偶然事件。

所以,如果没有理论抽象和机制解释,人类的因果推论就与动物界的条件反射无异了。休谟也看到,动物也能归纳经验现象的重复关联,但动物没有抽象思维能力,因此动物的“因果推论”一定是纯经验性的,由此可见人类因果推论的基础也是来自经验(Hume,1999/1748:165-168)。如果休谟看到巴甫洛夫的试验,也许他会重新考虑动物的“因果推论”与人类因果推论的区别。人类与动物思维的不同,就在于人类不满足于对恒定关联的简单归纳,人类能构造超越感官经验的理论模型,解释经验相关的内在机制。小耗子发现自己过去吃奶酪的地方没有奶酪了,会立刻换一个地方觅食,也不会精神崩溃,因为它们从来不会把过去的经验推论到未来之必然。只有人类才会因为自己的理论出现错误而精神崩溃。

人类的理论抽象一定基于经验观察,但从经验观察上升到理论的过程绝非简单归纳——比如从观察几只白天鹅到判断所有天鹅都是白的,或者听到铃声就想到食物的归纳——而是思维的跳跃,如从苹果落地到万有引力、恒定光速到相对论、企业存在到交易成本般的跳跃。无论是传统的神学和玄学理论还是现代的科学理论,都不是对经验观察的简单归纳。爱因斯坦提出他的能量公式 $E = MC^2$ 时,并没有任何经验资料,完全是在光电磁理论基础上的天才猜想。如果理论只是简单归纳,那么因果就真如休谟所言,是经验关联的代名词。

理论思维与经验归纳至少有几方面的区别:其一,理论自身的积累。理论抽象一定以现有理论为基础,要么是现有理论的应用,要么是对现有理论的革命或创新;即使创新,也往往用现有理论作为建筑材料,即所谓的批判继承。其二,普遍性。综合各种经验现象,而不是局限于具体经验现象。其三,理论自身必须有逻辑连贯性,自圆其说,不能自相矛盾。而经验世界却是充斥着矛盾的。

理论与因果机制的关系可从两方面来看。一方面,机制解释是理论解释。科学家对机制的定义是机械的,类似黑匣子里的齿轮和传动部件,由原因产生结果(Woodward,2002;Pearl,2000)。爱因斯坦比喻物理世界为“一个永远打不开的密封手表”,人类只能猜测里面的运作机制(Einstein & Infeld,1938)。社会科学家对机制没有统一定义,但似乎一致认为,社会机制是具备定律性质的理论概念(Hedstrom & Ylikoski,2010;Gross,2009;Hedstrom,2005;Kiser & Hechter,1991;Elster,1999)。具体经验现象千变万化,抽象的因果机制至少是相对稳定的。

理论家们争论社会机制是宏观的还是微观的,可观察的还是不可观察的(Gross 2009)。目前比较一致的观点认为社会机制应该具备以下三个特点:(1)时间过程;(2)集体或者个体行动者;(3)行动者对文化意义的定义和解释。而是集体还是个体层面的选择则取决于研究现象。比如解释国家行为需要分析政党、阶级、社团等;解释集体行为的逻辑则要从个人出发;而解释个人行为则需要从社会心理的角度,即个体对文化规范在认知和情感各方面的反应、对社会情景的解释、社会现实的理解,等等。所以,至少从个体角度来看,社会机制是难以直接观察的,是理论抽象(Hedstrom 2005)。

另一方面,并非所有理论都是关于因果机制的理论。理论可以是关于因果机制的理论,也可以是与因果分析无关的理论,还可以是超越具体因果机制的一般大理论。机制理论如文化再生产理论,解释父代教育如何影响下一代(Bourdieu, 1984/1979);内卷化理论,解释劳动力过剩如何导致农业劳动生产率下降(黄宗智 2006)。非因果理论如进化论的“人类是猴子变来的”这一命题,不能用相关验证,而是要找猿人和人猿的化石来验证。一般大理论如交易费用理论,解释企业组织形态,合同形式的选择,公地困境等。又如万有引力理论,解释苹果落地,潮汐变化,行星轨道,恒星衰变等等。并且,一般和具体是相对的,如量子力学还要问:万有引力的机制是什么?相关的发现和机制解释往往是两个不同过程,经常是发现相关在先,理论解释滞后,当然也有理论在先,经验验证滞后。例如,生物学家早就发现一个有趣的现象:各类生物的新陈代谢率与体积有相当固定的统计关系,从单细胞微生物到大型哺乳类动物(如鲸鱼、大象),新陈代谢率(Y)与体重(M)关系大约可以表述为 $Y = M^{3/4}$;心跳率(H)与体重的关系是 $H = M^{-1/4}$ 。简单说来,动物的体重增加一倍,能量消耗大概增加 $3/4$ 。科学家很长时间对这一比率感到困惑。多年之后,科学家才试图用生物新陈代谢网络(如血管网)的立体构造来解释这一现象(West & Brown 2005)。医药公司试验新药时,大概已经有新药的药理理论了。科学家在实验室查出艾滋病的元凶是HIV病毒后,还需要搞明白HIV致病的生物化学过程,这有赖生化理论的发展。中世纪欧洲黑死病肆虐,伦敦的一位医生发现黑死病与鼠患有关,消灭了老鼠就控制了黑死病。但以当时的医学理论,他无论如何也不可能搞明白,老鼠不过是病因的载体而已,真正病因是他肉眼看不到的细菌。不过从实用的角度来看,消灭了

老鼠,就切断了病菌的传播链。

社会现象比自然现象更复杂,有更大随机性,但是也有因果规律。社会科学的理论建设也应该寻找因果机制,并且要寻找具有普遍意义的因果机制(Merton,1968;Hedstrom,2005)。虽然社会学尚没有可以成为范式的统一大理论,但不乏机制理论。赫兹铎姆、斯韦勃格(Hedstrom & Swedberg,1996)用同一种因果机制,统一解释了三个著名社会学理论:默顿的自我实现的预言、科尔曼网络扩散理论和格兰诺维特的集体行为临界值理论。

自我实现预言是一种常见的社会现象,指最初不符合事实的预言(谣言),如果有足够的人相信,就会变假成真。比如一家资金状况良好的银行,有谣言说其面临倒闭,造成恐慌,引起越来越多储户挤兑,造成银行资金周转困难而倒闭。这里的机制是第一批挤兑的人给其他不明真相的人一个信号,增强谣言的可信性(Merton,1948)。

网络扩散理论是指一个新观念或新产品——如新药——其最初的传播往往依赖人际网络。例如人际网络在新药的最初传播过程中能起重要作用,是因为人们对新药的价值并无确定信息,所以更加依赖他人的判断。这时,每个人的行动会取决于其人际网络中他人的行动。当不确定性消失后,网络的作用亦消失(Coleman et al.,1957)。

集体行动的临界值指一个人决定是否参与一项集体行动,取决于已经参与该项集体行动的人数。不同人有不同的临界值。因而一项集体行动只需要一批发起人和一些临界值较低的个体,就可以在整个群体中传播开来。这种机制也许我们都有体验:当我们到了一个陌生的城市选择餐馆时,我们根据餐馆里已有客人的数目作出判断——没有客人的餐馆大概不会是好餐馆(Granovetter,1978)。

显而易见,这三类现象有共同因果机制,即在信息或制度不确定的情况下,个人的判断和行为取决于周围其他人的判断和行为。

四、因果理论的验证

如果理论的产生过程不是简单经验归纳而是飞跃,那么,我们如何跨越理论与经验之间的鸿沟?目前科学界比较普遍接受波普尔的证伪主义,又称后实证主义。波普尔强调:(1)演绎逻辑才是科学逻辑;(2)科

学理论是可以被证伪的一般性命题。就是说,科学发现不是从具体经验命题归纳出一般理论命题,而是从一般理论假设演绎出具体经验命题,后者可用经验资料验证(Popper,1959)。用演绎证伪代替归纳实证作为科学命题的逻辑,波普尔避免了经验实证主义的归纳困境。

所有关于事实的理论都是对经验关联的概括和解释,但并非都是科学理论。科学与非科学的区别不在于是否来源于经验,也不在于是否放之四海而皆准,而在于能否接受经验证伪。能被证伪,乃科学的本质(Popper,1959,1963)。比如传统中医学也源于经验积累,可是中医理论的阴阳五行解释套用《易经》辩证逻辑,概念(如寒、热、上火等)无法操作,理论解释(如冷热相克)无法证伪,因而也难以发展。

社会理论可以分为三类。第一类是不可验证的大理论,如历史唯物论、结构功能主义、制度经济学、理性选择理论等。第二类是可以经验验证的中层理论,如人力资本理论、社会资本理论、集体行动理论、阶级分析等等。第三类是经验理论,纯粹描述、概括经验资料,如升迁模型、流动表分析、小世界理论等等。第一类不能被经验验证,因而不属于科学理论的范畴。第三类“理论”也不属于波普尔意义上的理论命题,因为是纯经验性的。现代社会学忌讳普遍性命题,经验命题也可堂而皇之称为理论。中层理论属于科学理论范畴,一方面可以验证,另一方面又超越经验,具有普遍性(Merton,1968)。

第一、二类理论都能提供因果机制的解释。第三类“理论”只是对经验特质的描述或经验相关的概括,无因果机制解释。本文所谓因果理论,指关于因果机制的中层理论,不包括“天鹅都白、乌鸦都黑”,或者“社会主义都是香饽饽、资本主义都是臭豆腐”之类命题——此类命题能被单个个案证伪;也不包括“父亲教育影响儿子教育”,或“刘宗敏霸占了吴三桂的二奶”之类纯粹经验命题。

因果理论的验证,基本过程分三步:先从理论演绎因果命题,进而演绎相关假设,然后接受经验资料的检验。比如,假定理论T能说明A与B之间的因果机制。那么,若理论T正确,则A导致B;若A导致B,则A与B经验相关。如果统计分析证明A与B无相关,理论T被证伪;如果统计分析证明A与B有相关,理论T得到一次验证(corroborated)。这个过程可称理论检验三步曲,示于图2。

我将理论检验过程分为三个逻辑步骤是为了突出相关与因果的区别。如果有实验数据,无需顾虑伪相关和内生性问题,那么可以跳

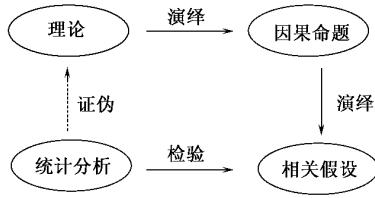


图2 检验因果理论的演绎逻辑

过相关假设,直接检验因果命题,这要比观察资料的检验更严格。但无论实验还是统计,都只能证伪不能证明理论 T,因为可能有其他理论也能解释 A 与 B 之间的因果机制。我们只能说理论 T 经过了一次经验验证。一般性理论只能接受一次又一次经验验证,但永远可能被新经验证伪。

中层因果理论除图 2 所示基本检验外,还有两种办法进一步佐证:一种可称为因果链检验,即测量因果机制中的某些可测量的变量;另一种可权且称为交叉检验,即用不同经验现象验证同一因果理论。

测量因果链条中的一环,即中间变量,是一种常用的验证因果理论的手段。中间变量虽然不是因果机制,但是因果链上的一环,证明中间变量的作用;虽然它不直接证明因果机制,但是一种间接验证。例如美国社会学界 20 世纪 60、70 年代十分流行的升迁路径模型,是一纯粹经验模型(Blau & Duncan, 1967)。核心路径是父代教育影响子代教育,子代教育又影响其职业。子代教育与职业的关系可以用人力资本理论解释。但是,父代教育与子代教育之间的相关仍然需要理论解释。后来法国社会学家布迪厄(Bourdieu, 1984/1979)提出了一个文化再生产理论,立刻引起了一向对理论、特别是外国理论不感兴趣的美国分层学者的浓厚兴趣,因为文化再生产概念为代际教育传递提供了理论解释。验证这一理论解释,可以通过在父代教育和子代教育之间加上文化资本这个中间变量来操作。虽然文化资本不是文化再生产,但的确是再生产过程中的一个重要环节……如此等等,可以细化我们对代际职业地位传递的理解。但是做定量研究的学者都有体会,中间变量往往不能完全解释原初变量之间的相关,例如升迁模型中控制了子代教育程度后,父代教育对子代职业仍有直接的作用;并且,中间变量越多,对原

初变量之间相关的解释力越弱。这可能是由于理论解释本身不完备,也可能是由于对中间变量的测量不精确造成的。

交叉检验指用不同经验现象验证同一因果理论。关于因果机制的理论应该是具有普遍性,应该能够解释不同的现象。下面以三项研究实践说明这种检验方法。这三项研究实践——笔者关于宗族与私营企业发展研究(Peng 2004)、宗族与生育率的研究(Peng 2010)和蔡晓丽关于宗族与政府问责的研究(Tsai 2007),虽然关注的是不同现象,但涉及相同的因果机制,有类似的理论解释。

在第一项研究中,笔者(Peng 2004)分析中国社会科学院《中国百县调查》365个村庄资料发现,在相等条件下,宗族聚居村庄的私营企业发展远远超过杂居村庄。援引科尔曼的社会资本理论和诺斯的制度经济学理论,笔者提出宗族网络以其凝聚性和信任,提高非正式规范的约束力;而在改革时期,正式产权制度缺位,宗族网络起到执行非正式产权(即农民的私有观念)、保护私营企业家利益的作用。所以宗族网络强的村庄,私营企业发展迅速。

传统网络分析学者对这一套新制度主义解释不以为然,特别是对宗族网络的规范约束力表示怀疑,他们的解释是:宗族网络能为族人带来资源,因而刺激私营企业发展。为了验证宗族的规范约束力,笔者又分析了宗族与生育率的关系。宗族的核心价值是生儿育女、延续香火。如果宗族网络有规范约束力,那么应该能够强化族人的生育观念,并且以其凝聚力协助族人抵制计划生育政策,超生超育。《百县调查》村资料的确表明,宗族网络能够提高生育率,虽然作用不是很大(Peng, 2010)。宗族网络的规范控制能力得到交叉验证。

蔡晓丽(Tsai 2007)是麻省理工学院政治系教授,研究正式与非正式制度对基层治理的作用。她结合田野调查和统计分析,发现宗族和庙会村庄公共物品(道路、学校、自来水等)的提供有积极作用。援引普特南的社会资本理论,蔡晓丽提出宗族网络能够提高村民的凝聚力和集体行动能力,使村民能够更有效地监督村政府履行其职责。该研究佐证了宗族的集体行动能力。规范控制是二级集体行动问题。

虽然这三项研究都不直接证明宗族网络的规范功能,但是三项独立的验证大大加强了这一理论解释的可信度。从联合概率的角度来看,这三项研究是偶合的可能性很小。

五、三点定位之我见：理论、定性和定量的整合

从方法论上来讲，定量、定性研究各有利弊，各有千秋。实证主义哲学强调假设检验，定量研究在社会科学中的应用偏向假设检验，又不关心假设从何而来，许多定量研究退化到纯粹描述相关或回归系数的归纳逻辑。但很显然，没有好的假设，检验也失去本意。个案研究不受定量研究中变量导向的制约，允许我们细致观察事件过程，分析行为动机，结合理论分析，探索因果机制，并提出假设。近几年社会科学中定量、定性两个阵营的敌对情绪逐渐式微，而结合两种方法的呼声越来越高，其中以三点定位原则最为有吸引力（King et al. ,1994；应用范例见 Uzzi ,1999；Tsai 2007）。

从语义上来说，三点定位（triangulation）来自几何学，即在立体空间中的任何三个点，都能准确定位空间中的任何其他点。比如全球定位系统，能根据三个不同位置的卫星，测定飞机或汽车的准确方位。运用于社会科学方法论，三点定位原则指从理论、个案研究和统计分析三个不同角度来论证研究者的假设。在我看来，三点定位的最大优势是整合了假设的产生和假设的检验过程。

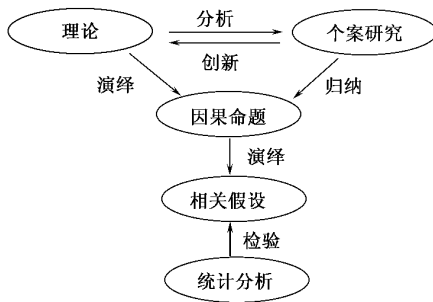


图3 因果分析中的理论、定性、定量三点定位

图3 勾勒理论、定性、定量三点定位的逻辑结构。这里所谓理论指有关因果机制的理论，如社会关系有助于克服集体行动困境，或者市场交易是有成本的等等。这里的所谓假设仅指关于经验相关的命题，如宗族网络能提高村庄公共物品的提供，或者乡镇企业比国有企业效率

高等等。假设是理论与经验证据之间的桥梁。

三点定位原则反映科学研究的完整过程,即从经验到理论,又从理论到经验。从研究的八股结构来看(彭玉生,2010),个案研究应处于理论和假设之间,辅助从理论到假设的演绎过程,或主导从个案到假设的归纳过程。在理想的研究过程中,假设来自现有理论、或个案观察、或二者的结合。

理论与个案的结合有几种不同情形。一个极端的情形是,现有理论能够精确解释我们所研究的现象,此类研究要么是对现有理论的验证,要么是应用性研究。若是验证理论,我们只需要从理论直接演绎出假设,个案研究的意义不大。另一个极端是,现有理论完全不适合新的经验现实,要被抛弃。这时个案研究的作用是归纳新概念、新理论,或者至少为理论创新提供方向。这种假设产生过程,我称之为个案主导理论归纳。

大部分研究者面对的是中间情况,即现有理论不能为经验问题提供精确答案,要么仅仅间接有关,要么本身存在谬误。这时,从理论到假设的逻辑演绎过程有许多环节,许多假定条件,如果没有实地观察和个案资料,我们很难提出具体假设以供统计检验。借助个案分析,我们就能理清这些环节,甚至适当修正理论。这种假设产生过程,我称之为个案辅助理论演绎。

总之,无论用主导理论归纳,还是个案辅助理论演绎,个案研究都能帮助我们提出事先假设。如此产生出来的假设再用统计资料严格证明,就完成了科学研究的一个完整循环。

笔者之所以强调假设要从理论中演绎产生,是为了避免循环论证。定量研究实践中,假设常常是统计结果的事后推论,理论是事后解释。比如事先看了统计结果,从中总结出理论解释,然后再反过来用统计资料验证理论。这显然有循环论证之嫌。事先假设之所以重要,是因为从错误理论出发,找到与之一致的统计资料的可能性很小,但从错误的统计结论出发(数据或模型失误)找到错误理论解释的可能性就大多了。即使统计结论正确,理论解释仍然可能错误。

如果理论解释来自现有理论,循环论证问题会被一定程度的减缓,因为现有理论是独立于我们手中的经验资料的。但是社会科学往往有多个理论范式和视角,事后理论解释仍然带有主观任意性。在这类情形下,实地观察和个案分析有助于我们从不同理论中做出正确的选择,

从而减低理论解释的任意性。所以,如果假设是统计结果的事后推论,个案与理论的结合能提高我们对结论的信心,但并不完成从经验到理论,又从理论到经验的完整过程。

如果事后理论解释是全新的,个案分析也能增加理论解释的可信度,但并不能验证之,因为错误理论总能找到与之一致的个案资料。个案研究本身就是理论与经验的逻辑循环。研究者能否通过个案分析发现理论错误,做正确理论解释,取决于研究者的观察能力和逻辑思维能力,并无程序可循。所以,事后理论解释无论多完美,仍然是假设,需要新定量数据的验证。

六、总 结

万物皆有联系。如何认识这种联系却成了哲学家争论不休的难题。因果联系是科学研究和日常生活中最常见的,而人类的感官并不能直接观察到因果作用,只能从重复的经验关联推论因果关系。如何从有限的经验归纳推论必然的因果关系,这就是休谟的因果归纳问题。

从认识论角度看,归纳问题的实质是人类有限感观和普遍理论之间的矛盾。如果放弃对一般理论的追求,满足于对现有经验的描述,归纳问题也就不存在了。但是,人类不满足于经验相关描述,如苹果熟了落地、耗子传染黑死病、社会关系影响集体行动等。人类是创造理论的动物。人类创造各种理论,解释经验关联的机制。并且,创造理论的过程决不是简单归纳,而是大脑的飞跃,爱因斯坦称之为“自由发明”,波普尔称之为“大胆假设”。大胆程度大概取决于研究者的理论想象力,所谓“天高任鸟飞,海阔凭鱼跃”。阴阳五行、风水八卦、上帝创世,当然还有科学理论大概都是这么产生的。但是飞跃的逻辑是未知的、甚至神秘的。我们只知道理论飞跃是基于经验资料分析和对现有理论的批判继承。除此之外,所知甚少。波普尔一整本书《假设与猜想》也没有说明大胆假设的逻辑规则。科学飞跃的过程与神学飞跃的过程是否有所不同,亦不得而知。

波普尔强调,科学理论与非科学理论的分水岭不在于是否来自经验,也不在于是否放之四海而皆准,而在于是否能够证伪。科学理论,通过演绎逻辑,推导出具体命题,接受经验资料验证。证伪的过程有比

较严格的逻辑和科学界公认的程序可循,不可任意跳跃。所以这是“小心求证”的过程(Popper,1963)。没有证伪,就没有理论发展。

从方法论上来看,密尔逻辑是对经验归纳的严格定义。密尔逻辑保留了休谟因果定义中的必然性,但指出了必然性因果推论的条件,即其他因素不变。只有控制了其他因素,我们才能从错综复杂、纷乱的具体经验现象中隔离、抽象出简单的、一般性的因果规律。只有控制其他因素,我们才能确保经验观察的可重复性,从有限经验观察,归纳出普遍结论。随机化试验设计使密尔逻辑具备可操作性,是因果推论的“金本位”。对于非实验数据,多元统计模型用变量控制近似试验控制,运用得当,也能比较有效地证明相关,并检验因果理论。

个案比较方法在社会科学中应用十分广泛,但不满足密尔逻辑的基本条件,因而不能从经验上论证因果。许多质性研究转向事件过程的叙述,即利用现有理论,从因果机制上解释具体的事件发生链。对具体因果链的解释,是理论演绎和经验分析的反复碰撞。从这个意义上讲,个案研究都是事后分析,可以提出假设而不能检验假设。事实上,许多质性研究者也拒绝使用假设检验之类的科学主义语言。

所以,理论、定性和定量各有自身的逻辑和优势,三者结合才能相得益彰。从经验观察到理论,再从理论到经验检验是认识过程的两个步骤,缺一不可。理论结合个案分析提出假设,再以统计数据检验之,有助于避免对统计结论的事后解释,这是三点定位原则的最重要功用。

实践中,未必三方面论证都在同一项研究中完成。工于理论抽象的大师能从理论到理论,概念到概念,提出新理论新概念,不屑于经验研究之琐累。富于洞察力的质性研究者专注于从个案观察或民族志研究中提出新理论,将定量验证的工作拱手相让。而善于统计分析的学者,未必对理论或个案有兴趣。当然还有时(如宏观历史研究)受资料所限,定量验证无用武之地,这时,理论论证便更为重要。

因果概念是人类的古老思维习惯,却又是一个难以捕捉的概念。历史上曾有不少人试图从人类思维中驱除因果概念(如大哲学家罗素),或者用相关的概念取而代之(如皮尔森),但是这个概念却有顽强的生命力,挥之不去。我认为因果命题和人类抽象理论紧密相连,企图摒弃因果概念,似乎徒劳无益。人类不能满足于对经验相关的描述和概括,而且还要解释经验相关的机制;机制解释就是理论解释。科学理论不断发展,人类对因果机制的解释也越来越深入。

参考文献:

- 黄宗智 2006,《长江三角洲小农家庭与乡村发展》,北京:中华书局。
- 2007,《连接经验与理论:建立中国的现代学术》,《开放时代》第4期。
- 彭玉生 2010,《“洋八股”与社会科学规范》,《社会学研究》第2期。
- 孙立平 2000,《“过程—事件分析”与中国农村中国家—农民关系的实践形态》,天则经济研究所演讲,1月28日(<http://www.unirule.org.cn/Secondweb/DWContent.asp?DWID=307>)。
- Abbott, Andrew 1995, “Sequence Analysis: New Methods for Old Ideas.” *Annual Review of Sociology* 21.
- 2001, *Time Matters: On Theory and Method*. Chicago: University of Chicago Press.
- Abell, Peter 1993, “Some Aspect of Narrative Method.” *Journal of Mathematical Sociology* 18(2-3).
- Beauchamp, Tom & Alexander Rosenberg 1981, *Hume and the Problem of Causation*. Oxford: Oxford University Press.
- Bendix, Reinhard 1956, *Work and Authority in Industry: Ideologies of Management in the Course of Industrialization*. Berkeley: University of California Press.
- Blalock, Hubert 1980, *Causal Inference in Nonexperimental Research*. New York: WW Norton & Company.
- Blau, Peter M. & Dudley O. Duncan 1967, *The American Occupational Structure*. New York: Free Press.
- Bonjour, Dorothe, Lynn F. Cherkas, Jonathan E. Haskel, Denise D. Hawkes & D. Tim 2003, “Returns to Education: Evidence from U. K. Twins.” *The American Economic Review* 93(5).
- Bourdieu, Pierre 1984/1979, *Distinction: A Social Critique of the Judgment of Taste*, (trans.) by Richard Nice. Cambridge: Harvard University Press.
- Coase, Ronald 1937, “The Nature of the Firm.” *Economica* 4(16).
- 1960, “The Problem of Social Cost.” *Journal of Law and Economics* 3.
- Coleman, James S., E. Katz & H. Menzel 1957, “The Diffusion of an Innovation Among Physicians.” *Sociometry* 20.
- Denzin, Norman K. & Yvonna S. Lincoln (eds.) 2000, *Handbook of Qualitative Research*. Thousand Oaks: Sage Publications.
- Einstein, Albert & Leopold Infeld 1938, *The Evolution of Physics*. New York: Simon & Schuster, Inc.
- Elster, Jon 1999, *Alchemies of the Mind: Rationality and the Emotions*. Cambridge, UK: Cambridge University Press.
- Fisher, Ronald A. 1935, *The Design of Experiments*. Edinburgh: Oliver and Boyd.
- Geertz, Clifford 1973, *The Interpretation of Cultures*. New York: Basic Books.
- Griffin, Larry 1993, “Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology.” *American Journal of Sociology* 98(5).
- Granovetter, Mark 1978, “Threshold Models of Collective Behavior.” *American Journal of*

- Sociology 83.
- 2007, “The Social Construction of Corruption.” In Victor Nee & Richard Swedberg (ed.) ,*On Capitalism*. California: Stanford University Press.
- Gross , Neil 2009, “A Pragmatist Theory of Social Mechanisms.” *American Sociological Review* 74.
- Hedstrom , Peter 2005 *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge , UK: Cambridge University Press.
- Hedstrom , Peter & Petri Ylikoski 2010, “Mechanisms in Social Sciences.” *Annual Review of Sociology* 36.
- Hedstrom , Peter & Richard Swedberg 1996, “Social Mechanims.” *Acta Sociologica* 39(3).
- Holland , Paul 1986, “Statistics and Causal Inference.” *Journal of American Statistical Association* 81 (396).
- Hume , David 2005/1739 *A Treatise of Human Nature*. New York: Barnes and Noble.
- 1999/1748 *An Enquiry Concerning Human Understanding*. Oxford , UK: Oxford University Press.
- Kant , Immanuel 1998/1781 *Critique of Pure Reason*. Cambridge: Cambridge University Press.
- King , Gary , Robert O. Keohane & Sidney Verba 1994 *Designing Social Inquiry: Scientific Inference in Qualitative Research*. New Jersey: Princeton University Press.
- Kiser , Edgar & Michael Hechter 1991, “The Role of General Theory in Comparative-Historical Sociology.” *American Journal of Sociology* 97.
- Kogut , Bruce & Udo Zander 2000, “Did Socialism Fail to Innovate? A Natural Experiment of the Two Zeiss Companies.” *American Sociological Review* 65.
- Kuhn , Thomas 1970 *The Structure of Scientific Revolution (2nd ed)*. Chicago: University of Chicago Press.
- Lee , Ching Kwan 1995, “Engendering the Worlds of Labor: Women Workers , Labor Markets , and Production Politics in the South China Economic Miracle.” *American Sociological Review* 60.
- Merton , Robert 1948, “The Self-fulfilling Prophecy.” *The Antioch Review* 8(2).
- 1968, “On Sociological Theories of the Middle Range.” In R. Merton , *Social Theory and Social Structure*. New York: The Free Press.
- Mill , John Stuart 2002/1848 *System of Logic: Ratiocinative and Inductive*. Honolulu , Hawaii: University Press of the Pacific.
- Miller , Paul , Charles Mulvey & Nick Martin 1995, “What Do Twins Studies Reveal About the Economic Returns to Education? A Comparison of Australian and U. S. Findings.” *The American Economic Review* 85(3).
- Moore , Barrington 1966 *Social Origins of Dictatorship and Democracy*. Boston , MA. : Beacon Press.
- Pearl , Judea 2000 *Causality: Models , Reasoning , and Inference*. London: Cambridge University Press.
- 2009, “Causal Inference in Statistics: An Overview.” *Statistical Surveys* 3.
- Pearson , Karl 1900 *The Grammer of Science (2nd Edition)*. London: Adam and Charles Black.

- Peng , Yusheng 2004, "Kinship Networks and Entrepreneurs in China's Transitional Economy." *American Journal of Sociology* 109.
- 2010, "When Formal Laws and Informal Norms Collide: Lineage Networks Versus Birth Control Policy in China." *American Journal of Sociolog* 116(3).
- Popper , Karl R. 1959 *The Logic of Scientific Discovery*. London: Routledge.
- 1963 *Conjectures and Refutations*. London: Routledge.
- Ragin , Charles 1987 *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.
- Rosenbaum , Paul R. 1984, "From Association to Causation in Observational Studies: The Role of Tests of Strongly Ignorable Treatment Assignment." *Journal of the American Statistical Association* 79(385).
- Russell , Bertrand 1903 *Principles of Mathematics*. London: Routledge.
- Skocpol , Theda 1979 *States and Social Revolutions: A Comparative Analysis of France , Russia , and China*. London: Cambridge University Press.
- Skocpol , Theda & Margaret Somers 1980, "The Uses of Comparative History in Macrosocial Inquiry." *Journal of Comparative History and Society* 22(2).
- Sobel , Michael 1995, "Causal Inference in the Social and Behavioral Sciences." In *Handbook of Statistical Modeling for the Social and Behavioral Sciences* ,(eds.) by G. Arminger , C. C. Clogg and M. E. Sobel. New York: Plenum Press.
- 1996, "An Introduction to Causal Inference." *Sociological Methods and Research* 24(3).
- Tsai , Lee Lily 2007, "Solidary Groups , Informal Accountability , and Local Public Goods Provision in Rural China." *American Political Science Review* 101(2).
- Time Asia 2000, "Taking on the System." *Time Asia*. October 9 , Vol. 156 , No. 14.
- Uzzi , Brian 1999, "Embeddedness in the Making of Financial Capital: How Social Relations and Networks Benefit Firm Seeking Financing." *American Sociological Review* 64.
- West , Geoffrey & James Brown 2005, "The Origin of Allometric Scaling Laws in Biology from Genomes to Ecosystems: Towards a Quantitative Unifying Theory of Biological Structure and Organization." *The Journal of Experimental Biology* 208.
- Winship , Christopher & Michael Sobel 2004, "Causal Inference in Sociological Studies." In *Handbook of Data Analysis* ,(eds.) by M. Hardy & A. Bryman. London: Sage Publications.
- Woodward , Jim 2002, "What is a Mechanism? A Counterfactual Account." *Philosophy of Science* 69.
- Zeitlin , Maurice 1988 *The Civil Wars in Chile: (or The Bourgeois Revolution That Never Were)*. New Jersey: Princeton University Press.
- Zhou , Xueguang 2005, "The Institutional Logic of Occupational Prestige Ranking: Reconceptualization and Reanalyses." *American Journal of Sociology* 111(1).

作者单位: 纽约城市大学

责任编辑: 张宛丽

PAPER

Causal Analyses in Social Sciences *Peng Yusheng* 1

Abstract: Hume’s problem of causal induction has inspired more than two centuries of epistemological discussion. J. S. Mill was the first one to elaborate the methodological principles of causal induction. The Millian principle of “overall similarities” for causal inference laid the foundation for Fisher’s randomized experimental design and guided causal analyses in multivariate statistical modeling and qualitative comparative case studies. Causal theories are not , however, induced from empirical correlations. They are produced through leaps of faith and empirically tested through deductive logic. The best research strategy is triangulation that combines theorizing , qualitative analysis , and statistical modeling.

On the Misusage and Adjustment of Household Members Matched Data : A discussion with the paper *Expansion of Higher Education and Inequality in Opportunity of Education* *Yang Ge & Wang Guangzhou* 33

Abstract: The data of household members matched are widely used in sociology , demography and related research , but the selection bias of the matched data is often ignored. This paper uses the raw data of the Fifth Census in 2000 and 1% Population Sample Survey in 2005 to match three kinds of relations , i. e. , father and sons , mother and children husband and wife , and confirms that the selection bias exist in these matched data grouped by age , gender , migration status , rural and urban distribution , education , geographical distribution , and so on. Based on the matched data , this paper re-tests the matched data , analysis models and conclusions of the paper *Expansion of Higher Education and Inequality in Opportunity of Education* and finds out that the selection bias of matched data would influence the accuracy of model analysis and research conclusion. For further reducing the impact of matched data bias , we propose two adjustment methods and find out that re-sampling and weighting methods can reduce the selective bias.

Selection Bias and Treatment Methods: A response to the questions