

政治学研究

比较案例研究中的或然性问题分析

叶成城 曹 航

(上海社会科学院 国际问题研究所, 上海 200020;
云南大学 “一带一路”研究院, 昆明 650091)

【摘要】比较案例研究是社会科学方法中最为核心的问题之一,密尔方法的应用是比较案例研究走向科学化重要标志。当前学界对于比较案例研究中密尔方法的认知仍存在困惑,即基于小样本的密尔方法何以得出比大样本的统计分析更为确定性的结论,因而,有必要澄清比较案例研究的或然性。无论是案例研究还是定量分析,都是或然性的,并且这两种传承的基础逻辑是一致的。比较案例研究的或然性源于社会科学与自然科学的差异,使得研究设计的各个步骤都产生确定性损失。具体而言,这种确定性损失来自于内部效度和外部效度:前者包括了对事实认定、寻找和检验机制、过程追踪与强化信念;后者包括了变量赋值、基于布尔代数的因果推断和理论适用范围推广,等等。因而,研究者在进行比较案例研究时,更加适合使用贝叶斯逻辑,将因果过程观察值和数据集观察值视为调整对理论主观信心的证据,而非在小样本的基础上去寻求确定性的解释。

【关键词】案例研究 或然性 内部效度 外部效度 确定性损失 贝叶斯逻辑

比较案例研究是社会科学方法中最为核心的问题之一,早期的社会科学家从纯粹的历史叙述开始过渡到具体的案例比较研究后,最为广泛采用的是英国哲学家约翰·密尔(John S. Mill)所提出的5种因果识别的逻辑方法,分别是求同法、求异法、求同求异法、共变法、求余法。^①密尔方法基于布尔代数逻辑,其中最为核心的是求异法。20世纪70年代以来,亚当·普沃斯基(Adam Przeworski)等人基于求异法提出了案例比较的“最相似系统设计”原则。^②求异法的核心逻辑是在其他条件一致时,导致结果不同的差异

即为原因。在如下两个案例中,如果因素A的差异导致了结果Y的差异,就可以认为A是Y的原因。

然而,研究者对于质性研究方法的“概率性”上仍然存在困惑:求异法基于最大相似性所使用的案例是表明了A和Y之间的必然性联系,还是说仅仅提高了关于“A是Y的原因”这一因果推断的可信度?

$$A+B+C+D+E \rightarrow Y$$

$$\sim A+B+C+D+E \rightarrow \sim Y$$

传统的充分性和必要性的分析往往强调前者,但是随着定性和定量之间争论的增加,当我们讨论“两

[作者简介] 叶成城,上海社会科学院国际问题研究所副研究员。

曹 航,云南大学“一带一路”研究院、国际关系研究院助理研究员。

◎ 本文系国家社科基金后期资助项目“早期西欧国家的构建与兴衰(1492—1848)”(项目批准号:21FGJB006)的阶段性成果。

① John S. Mill, *System of Logic: Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation* (New York: Harper & Brothers Publishers, 1898) 478-537.

② Adam Przeworski and Henry Teune, *The Logic of Comparative Social Inquiry* (New York: Wiley-Interscience, 1970) 31-35.

种传承”的因果性时,却会出现相应的逻辑困境与迷思。^① 首先,基于密尔逻辑的小样本研究只讨论变量同结果之间的必然性问题,往往涉及到结果出现与否的充分性、必要性问题。其中一个经典的模型是双变量模型是变量(A * B)共同组成了结果(Y)出现充分非必要条件,而变量A和B被视为Y的INUS条件。^② 其次,基于中等样本的定性比较分析(Qualitative Comparative Analysis, QCA)时常是被认为介于定性和定量之间的方法。QCA则是给出一个近似于充分/必要条件的概率分析(它也因此被认为更加接近于定量分析)。^③ QCA也存在不同类型,即便传统的清晰集也同样需要认为设定充分条件的标准,往往采用90%和95%作为充分条件来进行稳健性检验。^④ 而模糊集则在此基础强调变量对于集合的模糊属性,即变量的赋值不再是0和1的二分,而是更广泛具有(0~1)中间地带的赋值属性。^⑤ 尽管QCA的结果可以是概率性的,但是如果案例完全契合理论预期(在部分中等研究中也确实会出现),仍然会有一致性或覆盖率为1的情况。再次,基于大样本的因果推断则是根据统计方法得出一个概率性分析。其中最为经典的著作当属加里·金(Gary King)、罗伯特·基欧汉(Robert Keohane)和悉尼·维巴(Sidney Verba)所合著的《社会科学中的研究设计:定性研究中的科学推论》(Designing Social Inquiry: Scientific Inference in Qualitative Research),三位作者(以下简称为KKV)倡导以因果推断为导向的社会科学研究,察看自变量(X)对于结果出现(Y₁)或不出现(Y₀)概率所产生的影响。^⑥ 显然,回归分析都是概率性的,几乎很少出现P值为1的情况。

这里可以看到,案例研究的逻辑似乎出现了悖

论:为何少小样本反而可以比大样本研究得出更具确定性和必然性的结论?同时,一些研究也开始质疑混合方法运用的有效性,即认为逻辑的不一致使得上述方法共同使用是不恰当的,例如在小样本研究中提出有关充分/必要条件的因果框架,却用金等人的建议来指导进行研究设计。^⑦ 本文的目的在于解决比较案例研究中这一看似矛盾的问题,展示或然性在此类研究中的本体论和认识论来源,澄清比较案例研究的或然性。本文认为,无论是案例研究还是定量分析,都会产生或然性,这也意味着所谓的“两种传承”背后其实共享着同一种基础逻辑。

文章以下内容包括四个部分:第一部分将文献分为了定性、定量和混合方法三类,分析国内外研究对此讨论的不足之处。为了便于后续讨论,第二部分从案例的内外效度和理论的内外效度两个视角澄清和分析了效度的不同类型。第三部分从案例的内部效度、案例的外部效度和理论的外部效度视角出发,讨论了对变量与事件的认知、理论构建、确定机制与增强理论信念、适用性的进一步推广等阶段如何产生确定性损失,从而分析了案例研究的或然性来源。最后是简短的结语。

一、既有文献及其不足之处

因果解释是社会科学中案例研究的根本目的之一,对于因果解释的倡导可以追溯到亚里士多德和大卫·休谟(David Hume)。亚里士多德提出了基于质料、形式、动力和目的四种原因来解释事物。^⑧ 在统计技术尚未出现之前,前社会科学时代的先驱主要是以定性方法为主,他们已经隐约意识到了案例研究中的或然性。例如休谟在提出了关于因果识别的时间性、

① James Mahoney, "After KKV: The New Methodology of Qualitative Research," *World Politics* 62.1 (2010): 120-147; Gary Goertz and James Mahoney, *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences* (Princeton: Princeton University Press, 2012).

② INUS是Insufficient but Necessary part of an Unnecessary but Sufficient condition的缩写,意思是“某个充分不必要条件中的必要不充分部分”。国内一些学者的具体介绍参见冉奎、陈刚:《论因果关系的INUS理论》,《自然辩证法研究》2015年第6期。

③ Charles Ragin, *The Comparative Method: Moving Beyond Qualitative and Quantitative Methods* (Berkeley: University of California, 1987); Charles Ragin, *Redesigning Social Inquiry: Fuzzy Sets and Beyond* (Chicago: University of Chicago Press, 2008).

④ 具体的做法参见唐睿、唐世平:《历史遗产与原苏东国家的民主转型——基于26个国家的模糊集与多值QCA的双重检测》,《世界经济与政治》2013年第2期。

⑤ Charles Ragin, *Fuzzy-Set Social Science* (Chicago: University of Chicago Press, 2000) 149.

⑥ 为了便于叙述,将作者简称为KKV,将该书简称为DSI。参见Gary King, Robert Keohane and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton: Princeton University Press, 1994).

⑦ 陈超、李响:《逻辑因果与量化相关:少案例比较方法的两种路径》,《公共管理评论》2019年第1期。

⑧ 亚里士多德著,张竹明译:《物理学》,北京:商务印书馆,1982年,第49~53页。

恒常联结、充分必要性、求同求异法等八个原则的同时,也阐述了因果解释存在或然性的原因,例如存在不完全经验、相反的原因发生、类比发生的精确性等等。^①密尔在提出五种逻辑方法时,也论及了这些方法的适用范围和可靠性问题,这其实就隐含了密尔方法进行因果解释所面临的或然性问题。^②20世纪70年代之后,随着各类研究方法被逐渐推广和普及,案例研究的或然性问题反而一定程度被忽视了。当前主要有三类方法的讨论观点涉及到了比较案例研究中的或然性问题。

第一类是定性研究方法。然而在定性方法被普及之后,上述讨论中隐含的概率性问题似乎被早期的定性研究学者所忽视,往往得出必然性结论。其中一种常见的表述是定性研究中的2×2表:在特定情境下,自变量 X_1 和自变量 X_2 分别为因变量 Y 出现的必要条件, X_1 和 X_2 同时出现时为 Y 出现的充分条件。早期多数的案例比较研究都是以此作为基本的理论模型。^③这部分研究的缺陷在于对小样本研究和集合论方法的认知存在不足,或者仅仅将案例内研究视为基于贝叶斯方法的隐喻,而没有在方法层面真正重视和具体化。^④换言之,早期关于定性方法的讨论都是以决定论为前提的,自然也无法回答前文提及的“自由度越低反而确定性越强”的疑惑。

第二类是定量研究方法。大样本研究方法试图

通过统计分析来进行因果推断,这类方法时常会否认小样本研究外部效度或科学性,因其变量过多而案例过少从而导致了较低的自由度。早期部分学者甚至否认质性研究中的个案研究和案例比较的科学性。^⑤KKV系统性地提出了基于因果推断的样本分析方法,并将这些方法运用于案例研究中,例如区分案例研究的系统过程和非系统过程,强调数据收集的无偏性、有效性和一致性。^⑥然而KKV将定量操作方法直接“移植”到定性研究中的做法,遭致了许多定性或混合方法倡导者的批评,因为许多基于定量的方法很难在定性研究中实现,因而也无法解释小样本研究中的或然性问题。

第三种观点源于倡导混合方法的学者。他们认为,定性和定量分析的逻辑是不同的,定性研究采取的是案例内分析,旨在寻找自变量在因果过程中产生影响的证据。^⑦詹姆斯·马奥尼(James Mahoney)认为,需要区分两种方法间的差异,定量研究试图寻找案例间的数据集观察值,而定性研究则是寻求因果过程观察值,即前者讨论原因的结果,而后者讨论结果的原因。^⑧乔纳森·谢弗(Jonathan Schaffer)认为,对因果关系性质的主要论述分为概率提升和过程联系的观点;在概率提升的观点中,因果关系源于原因存在/不存在时,结果出现的概率差异;在过程联系的

① 大卫·休谟著,关文译:《人性论》,北京:商务印书馆,2016年,第161~173、194~197页。

② John S. Mill, *System of Logic: Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation* (New York: Harper & Brothers Publishers, 1898) 478-537.

③ 国内外许多具体的实证研究案例都是参照这个模式。国外的研究参见西达·斯考切波著,何俊志等译:《国家与社会革命》,上海:上海人民出版社,2013年;布莱恩·唐宁著,赵信敏译:《军事革命与政治变革:近代早期欧洲的民主与专制之起源》,上海:复旦大学出版社,2015年;Daniel Ziblatt, *Structuring the State: The Formation of Italy and Germany and the Puzzle of Federalism* (Princeton: Princeton University Press, 2006); James Mahoney, *The Legacies of Liberalism: Path Dependence and Political Regimes in Central America* (Baltimore: Johns Hopkins University Press, 2001); Dan Slater, *Ordering Power: Contentious Politics and Authoritarian Leviathans in Southeast Asia* (New York: Cambridge University Press, 2010)。国内的研究参见包刚升著:《民主崩溃的政治学》,北京:商务印书馆,2014年;唐睿著:《体制性吸纳与东亚国家政治转型》,北京:中央编译出版社,2014年;叶成城:《第一波半现代化之“帝国的胎动”——18世纪普鲁士和奥地利的崛起之路》,《世界经济与政治》2017年第5期。

④ Timothy J. McKeown, “Case Studies and the Statistical Worldview: Review of King, Keohane, and Verba’s *Designing Social Inquiry*,” *International Organization* 53.1 (1999) 161-190; Derek Beach and Rasmus Brun Pedersen, *Process-Tracing Methods: Foundations and Guidelines* (Ann Arbor: University of Michigan, 2013)。

⑤ Alexander L. George and Timothy J. McKeown, “Case Studies and Theories of Organizational Decision Making,” *Advances in Information Processing in Organizations* 2.1 (1985): 21.

⑥ Gary King, Robert Keohane and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton: Princeton University Press, 1994)。

⑦ David Collier, James Mahoney and Jason Seawright, “Claiming Too Much: Warnings About Selection Bias,” eds. Henry E. Brady and David Collier, *Rethinking Social Inquiry: Diverse Tools, Shared Standards* (Lanham: Rowman & Littlefield Publishers, 2010) 95-98.

⑧ James Mahoney, “After KKV: The New Methodology of Qualitative Research,” *World Politics* 62.1 (2010): 120-147; Gary Goertz and James Mahoney, *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences* (Princeton: Princeton University Press, 2012)。

观点中,因果根源在于从原因到结果之间的过程链。^①上述方法仍然没有完整地讨论定性研究中准实验状态和案例分析中的或然性问题,早期的多数研究仅仅强调定性和定量方法的混合使用而非融合,^②仍然认为充分/必要性逻辑与概率学派逻辑存在本质性差异。^③因此涉及到混合方法时,仍然会出现内部效度与外部效度无法统一的情况,因为前者基于过程追踪得出必然性结论,而后者则基于对统计分析得出或然性结论。

基于过去讨论的不足,要统一两种传承的逻辑,更加适合从概率的视角,将具体事件中特定因素对结果的贡献视为一种因果“贡献度”,从而解决个案中的概率性问题。^④在具体分析中,还需对客观概率和主观概率进行区分,前者认为概率是两个事物之间客观的物理性联系,后者则认为概率是个体的主观信念。^⑤因而,本文将以上述讨论为基础,从或然性的视角分析小样本研究中的内外效度问题。

二、比较案例研究中内外部效度的来源

在分析案例研究的或然性之前,需要讨论案例研究的效度来源,以澄清相关概念。具体而言,研究存在两方面的效度来源,即内部效度和外部效度。约翰·格林(John Gerring)给出如下定义:内部效度指假设相对于样本(研究者实际研究的案例)的正确性,外部效度指的是假设相对于推断总体(未研究的案例)的正确性。^⑥上述定义的缺陷在于没有区分案例的效度和理论的效度,因此本节将从案例的效度和理论的效度两个维度来厘清效度的概念。

1. 案例的效度

对于案例研究而言,其内部效度指的是理论在解释个案本身时的有效性,即对个案的过程追踪是否有效支持理论假设,而在具体研究中的外部效度则指的是在具体研究所限定的时空情境之内,案例研究对于

总体样本的有效性。在设定时空情境后,样本数量会大幅减少,尤其诸多以国家为单位的研究,时常只有少量样本,但是这些小样本却已经是总体,即全样本研究。假设在时空情境 T 中,一共只有四个案例 C₁、C₂、C₃ 和 C₄,研究者试图用变量 A 和 B 来解释案例中结果 Y 的差异,这需要两方面的效度:一方面,研究者需要分别对这四个案例进行过程追踪来检验理论的有效性,即从内部效度的视角看理论是否符合预期;另一方面,从外部效度的视角看,研究者则是通过密尔方法来检验因素的变化是否符合其理论预期,基于类型学分析来解释关键变量的赋值变化同结果 Y 之间的相关性,即前文所述传统的双变量经典模型。

2. 理论的效度

理论的内部效度指的是其在所限定时空情境中对全部案例的有效性,而其外部效度则指的是其在所限定时空情境之外案例的有效性。沿用上文的例子,理论的内部效度相当于将在时空情境 T 中 C₁、C₂、C₃ 和 C₄ 内外效度的加总,也就是将时空情境限制下的全样本视为一个整体,研究理论在上述情境下的有效性。通常而言,小样本研究的理论内部效度并不存在太多争议,内部效度的质量取决于研究的内部结构和设计能否有效地支持研究结果。研究者可以通过更严格的条件控制,例如西德尼·塔罗(Sidney Tarrow)提出使用配对比较,基于最大相似性来选择案例以及解释变量之间的相关性,并同时两个案例进行过程跟踪,从而最大程度地接近“准实验状态”来强化理论的内部效度。^⑦

而外部效度考察理论在其他环境或不同样本中的推广性和泛化能力。小样本研究的理论外部效度往往存在较多的争议,因为很难保证超越情境之后的理论解释力。因为当情境存在较大差异时,即便相同机制也可能导致获得不同的结果。^⑧ 尽管通过小样本

① Jonathan Schaffer, "Causes as Probability Raisers of Processes," *The Journal of Philosophy* 98.2 (2001): 75-92.

② Evan S. Lieberman, "Nested Analysis as a Mixed-Method Strategy for Comparative Research," *American Political Science Review* 99.3 (2005): 435-452.

③ Ingo Rohlfing, "What You See and What You Get: Pitfalls and Principles of Nested Analysis in Comparative research," *Comparative Political Studies* 41.11 (2008) 1492-1514;陈超、李响:《逻辑因果与量化相关:少案例比较方法的两种路径》,《公共管理评论》2019年第1期。

④ 何朝安:《因果度问题》,《自然辩证法研究》2019年第7期。

⑤ Ellery Eells, *Probabilistic Causality* (New York: Cambridge University Press, 1991) 34-36.

⑥ John Gerring, *Case Study Research: Principles and Practices* (Cambridge: Cambridge University Press, 2006) 217.

⑦ Sidney Tarrow, "The Strategy of Paired Comparison: Toward a Theory of Practice," *Comparative Political Studies* 43.2 (2010): 230-259.

⑧ Tulia G. Falleti and Julia F. Lynch, "Context and Causal Mechanisms in Political Analysis," *Comparative Political Studies* 42.9 (2009): 1143-1166.

研究获得理论的外部效度具有难度,但也并非完全不可实现,具体来说可以包含如下两方面思路:第一种是研究者总结恒常联结的因果过程,即发现深度案例研究中紧密相连的因果链,当这些因果链不难通过逻辑演绎和现有知识进行理解时,其中的机制就具备了较强的外部效度。柯南道尔(Conan Doyle)曾借福尔摩斯之口道出了这种恒常联结的作用:“全部生活就是一个大链条,只需见其一环,就可知其整体与性质。”^①社会科学中的一个著名的案例是斯科特对于东南亚村庄农民“开小差”机制的深刻观察,尽管没有大样本统计分析,但是这种研究仍然有着强外部效度,因为这种“弱者的反抗”的行为模式或机制并非当地农民所特有的,而是人们普遍存在的日常行为。^②第二种来源则是基于内部效度的强理论化。丹·斯雷特(Dan Slater)和丹尼尔·齐勃拉特(Daniel Ziblatt)提出了控制比较(Controlled Comparison)的理念,主要通过三方面的策略提高理论的外部效度:用一般的变量或机制进行操作化,寻求代表性变化(variation)以试图反映更广泛的群体,并通过理论选择案例以最大限度地实现控制。^③上述三个策略有助于建立一个更具外部效度的理论,并通过寻找其他理论所无法解释的现象/证据来否定其他竞争性解释,从而强化自身的外部效度。一个经典的案例是《物种起源》,达尔文仅仅观察了数十种生物的演化过程,就总结出“变异—选择—遗传”的生物演化理论,其外部效度强到几乎不受任何条件限制。^④如果从量化分析的角度来看,数十种生物相对地球上物种总数而言完全不具备代表性,但人们都对其机制的有效性深信不疑。因为达尔文提供了大量证据,这些证据是创造论等其他竞争性解释所无法解释的。例如他观察到一些生物体中存在一些没有明确功能或目的的器官与结构,创造论中很难解释为何造物主会“如此无聊”地创造无意义的器官,但可以通过进化论解释为早期生物的功能遗留或新功能的演化。

在区分了基于案例的内外效度同基于理论的内外效度的差异之后,需要指出的是,本文所讨论的或然性仍然侧重于案例的内外效度或者说理论的内部效度,因为理论的推广所面临的环境更加复杂,时常需要具体问题具体分析,而案例的内外效度更适合标准化的操作流程。

三、比较案例研究中的效度与或然性的多重来源

多数文献对于小样本研究的内外效度仍然是基于充分/必要性的思维,这种充分/必要性思维在一些环境相对单一的自然科学中(如经典物理学)可能是成立的,而在系统和环境复杂的学科中(如社会科学),就会由于复杂的背景因素的干扰而产生诸多的确定性损失,使得研究者需要从概率的视角来理解因果关系。本节将从案例的内部效度、案例的外部效度和理论的外部效度来讨论比较案例研究的或然性来源。其中案例研究的内部效度更依赖于数据集观察值(Data-Set Observation, DSO),而它的外部效度更加依赖于因果过程观察值(Causal-Process Observation, CPO)。^⑤

1. 案例的内部效度与或然性来源

案例研究的内部效度主要源于CPO,因此其内部效度的或然性实际上是源于CPO的特征,这些特征使得在研究者观察因果关系的过程中,存在从本体论到认识论的确定性损失。马奥尼将CPO分为三种类型,分别是自变量CPO、机制CPO和辅助结果CPO,这三类CPO都存在其或然性来源。^⑥

首先是自变量因果过程观察值(Independent Variable CPOs),它主要用于衡量变量的赋值或存在性,即 $X=1$ 是不是真实存在的。对于变量赋值的一定程度涉及了部分科学哲学对于本体论和认识论层面关于实在性的讨论。罗伊·巴斯卡(Roy Bhaskar)将人们对于世界的了解区分为真实域(domain of

① 柯南道尔著,俞步凡译:《血字的研究》,上海:上海社会科学院出版社,2012年,第25页。

② 詹姆斯·斯科特著,郑广怀等译:《弱者的武器》,南京:译林出版社,2011年。

③ Dan Slater and Daniel Ziblatt, “The Enduring Indispensability of the Controlled Comparison,” *Comparative Political Studies* 46.10 (2013): 1301–1327.

④ 参见达尔文著,周建人等译:《物种起源》,北京:商务印书馆,2017年。

⑤ Henry E. Brady and David Collier (eds.), *Rethinking Social Inquiry: Diverse Tools, Shared Standards* (Lanham: Rowman & Littlefield Publishers, 2010) 277–283.

⑥ 三种观测值的具体区分参见 James Mahoney, “After KKV: The New Methodology of Qualitative Research,” *World Politics* 62.1 (2010): 120–147。

real)、实际域(domain of actual)和经验域(domain of empirical)。^① 经验域仅包含了日常的经验,实际域包含了经验观察之外的事件,真实域还包括超出直观感受的机制。^② 经验与事件的差异本身就是案例研究或然性的来源。经验域是日常生活所观察、体验到的事物,例如一个人在上海亲身经历下雨,或亲眼目睹了X枪杀Y,通常无须怀疑其真实性。而那些非亲历的事件就不再是经验,而属于实际域,只能通过相关的记载或证据去推测其真实存在,例如对在上海的人而言,北京下雨就只是通过资讯了解的“事件”,观察者只能通过各类证据无限接近于真实性,很难认为存在理论意义上100%的“完美的充分性”。用一个侦探小说的例子来说明,假设要讨论一个因果关系:嫌疑人X枪杀了Y,他是凶手。第一种情况是X枪杀Y的全过程被现场直播,那么这个因果过程是100%确信的。第二种情况则是上述过程是隐秘的,例如侦探只能间接地通过证据证明X为凶手。没有亲历的事件只能通过证据来接近其原本的真相,多数情况下这是概率问题,即既不存在绝对的必要性(例如不在场证明可能是伪造的),也不存在绝对的充分性(例如被人设计陷害)。即便这些概率很小,但是它可以使得所谓的“充分必要性”仅仅是接近于95%或99%,却无法达到100%。

其次,机制因果过程观察值(Mechanism CPOs)则涉及巴斯卡所提及的真实域。在真实域中,CPO的或然性更为明显,因为机制无法仅仅通过经验或者事件获得。马里奥·邦格(Mario Bunge)给机制下的定义是:机制是一个系统中一系列进程的集合,以至于在系统整体中这些过程带来或阻止某种变化,无论它是一种属性还是其他过程的涌现(emergence)。^③ 当研

究者试图去理解机制并以此来进行因果解释时,只能通过过程追踪和半负面案例等方式找到相应证据来部分接近真相。^④ 许多将寻找机制视作是过度决定论的观点,事实上是错误理解了机制的内涵,尝试寻求覆盖律解释(covering law explanation)或者类律规律性(law-like regularities)。^⑤ 当机制被理解为进程的集合时,就意味着本体论层面的机制远比认识论层面的复杂,研究者只能通过过程追踪和理论化将机制描绘为有限的步骤,例如将真实的机制: $Y_{(0)} \rightarrow Y_{(1)} \rightarrow Y_{(2)} \rightarrow \dots \rightarrow Y_{(n)}$ (真实的机制)简化为 $Y_{(0)} \rightarrow Y_{(m)} \rightarrow Y_{(n)}$ (一种机制的模型)。因而在主观层面认识客观机制的过程也不存在绝对的准确性,而只能不断地去接近真实的机制,理解机制仅是“强调因变量在因果机制各环节的具体变化过程”。^⑥

第三,对辅助结果的过程追踪观察值(Auxiliary Outcome CPOs)由于遵循了贝叶斯逻辑,则天然具备了或然性。马奥尼认为,辅助结果CPOs是独立的事件,如果理论假设成立,它们就应该产生,因而它可以被看作是由产生结果的原因所留下的额外“痕迹”或“标记”。^⑦ 当理论假设认为可能存在两种不同的竞争性机制A和B时,如果可以发现解释成立时所独有的“痕迹”或“标记”,就会大幅增加A的解释力并降低B的解释力。这在数学上则用贝叶斯定理来描绘: $P(h/e) = P(h) \times P(e/h) \div P(e)$ 。这里P(h)是假说的“先验概率”,P(e)是证据概率,P(e/h)是当假说在正确的条件下证据出现的概率,而P(h/e)则是在证据成立的条件下假说成立的概率。简而言之,人们会根据证据对于假说的支持程度来不断调整对理论的信心。例如在《银色马》的案例中,福尔摩斯因为当晚看门狗没有叫而增加对解释A(马夫监守自盗)的信

① Roy Bhaskar, *A Realist Theory of Science* (London: Verso, 2008) 56-62.

② Roy Bhaskar, *A Realist Theory of Science* (London: Verso, 2008) 1.

③ Mario Bunge, "Mechanism and Explanation," *Philosophy of the Social Sciences* 27.4 (1997): 410-465; 马里奥·邦格著,李宗荣译:《涌现与汇聚:新质的产生与知识的统一》,北京:人民出版社,2019年,第26页。

④ 叶成城、唐世平:《基于因果机制的案例选择》,《世界经济与政治》2019年第10期。

⑤ 这些问题源于乔恩·埃尔斯特(Jon Elster)等人对机制的误解,将机制理解为因果律的文献很多,例如Jon Elster, *Alchemies of the Mind: Rationality and the Emotions* (Cambridge: Cambridge University Press, 1999) 1; Stuart Glennan, "Mechanisms and the Nature of Causation," *Erkenntnis* 44.1 (1996): 52.

⑥ Derek Beach and Ingo Rohlfing, "Integrating Cross-Case Analyses and Process Tracing in Set-Theoretic Research: Strategies and Parameters of Debate," *Sociological Methods & Research* 47.1 (2018): 12-15.

⑦ James Mahoney, "After KKV: The New Methodology of Qualitative Research," *World Politics* 62.1 (2010): 129-130.

心,并降低了对解释 B(外人入侵)的信心。^①

需要指出的是,这些因果过程观察值是基于因果机制和贝叶斯逻辑层面的充分/必要性,它的本体论涵义是不同于基于因果律和演绎—法理模型(deductive-nomological model)中的充分/必要性的。^②在演绎—法理模型中,例如存在命题“天鹅是白色的”,就意味着白色是天鹅的必要条件,一旦观察到黑天鹅就会完全否定上述命题。^③而过程追踪和贝叶斯逻辑则是一种归纳模型(inductive model),通常是借助证据点达到相对足够的可信度就可以认为接近必要或充分条件。类似于法官或陪审团的断案,如果有着相对足够的证据(例如 90%的主观概率),就倾向于让这些证据成为被告人有罪的“充分条件”,而通常不会依靠演绎法或者等待 100%的绝对证据。

2. 案例的外部效度与或然性来源

案例的外部效度来自于数据集观察值,通过样本之间的比较来进行因果推断,从而获得多个案例之间的理论解释力。然而,由于本体论和认识论之间的缺口以及社会科学和自然科学的差异,案例的外部效度同样存在或然性。

一方面,对多案例的赋值比单案例的赋值要更加困难,跨案例比较研究中的赋值过程时常包含或然性。社会科学家感兴趣的许多事物并不能整齐地归入清晰的集合,这就意味着用 0 和 1 进行区分的清晰集来分析社会现象存在诸多困惑,例如存在许多介于民主和非民主国家之间的案例,美国在发达国家中赋值可以为 1,但在民主国家中赋值可能就是 0.9。^④因此,在关于民主和平论的检验中,就需要权衡该理论只适用于民主程度更高的国家,还是全部民主国家。因为当存在介于隶属于/不隶属于该集合时,就需要模糊集方法和使用或然性分析来区分其中的“充分性”或“必要性”。^⑤如表 1 所示,以民主和平论为例,假设控制变量 A 和因变量 Y(是否发生战争)都是二

分变量,而主要的解释变量 B(民主程度)则存在不同程度的定义。可以看到,案例 1 和案例 5 是典型的求异法,由此推断变量 B 是结果 Y 的原因。但存在案例 2、案例 3 和案例 4 时,就需要思考更多或然性问题,即变量 B 处于何种程度会对结果的出现产生影响,从而只能够得出结论认为民主程度对于战争产生了负面的促进作用而无法认定为充分条件。

表 1 模糊集与变量赋值的或然性来源

案例	变量 A	变量 B	因变量 Y
案例 1	1	1	1
案例 2	1	0.75	1
案例 3	1	0.5	0
案例 4	1	0.25	0
案例 5	1	0	0

注:表 1 为笔者自制。

另一方面,案例比较研究主要通过密尔方法和布尔逻辑来对数据集观察值进行因果推断,数据集观察值外部效度的或然性源于密尔方法本身的局限性。理解密尔方法要从求异法出发,因为求异法是其其他方法的基础。^⑥求异法的有效性源于其强有力的反事实逻辑,即通过控制当原因出现/不出现时,结果出现明显差异。求异法试图通过单一变量的变化来获得因果解释,即它的充分性前提是其其他条件不变(Ceteris Paribus),但世界上几乎很难在非实验状态下找到一对只有一个差异而其他完全相同的样本。因此,对于求异法而言,其确定性只存于严格实验状态下、层次结构相对单一的学科,例如物理学、化学和部分生物学领域,而在多数社会科学领域,都很难依靠严格意义上的实验状态来获得绝对的确定性。与自然科学不同的是,在社会科学领域中能够通过实验来解答的问题不仅少之又少,而且能通过控制实验来解答的问题往往只是随处可见和用其他

① 《银色马》与过程追踪的方法论探讨参见 David Collier, “Understanding Process Tracing,” *PS: Political Science & Politics* 44.4 (2011): 823–830; James Mahoney, “The Logic of Process Tracing Tests in the Social Sciences,” *Sociological Methods & Research* 41.4 (2012): 574–583.

② James Mahoney, “The Logic of Process Tracing Tests in the Social Sciences,” *Sociological Methods & Research* 41.4 (2012): 583–584.

③ Carl Hempel, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: Free Press, 1965) 26.

④ Charles Ragin, *Redesigning Social Inquiry: Fuzzy Sets and Beyond* (Chicago: University of Chicago Press, 2008) 29–30.

⑤ 参见 Charles Ragin, *Fuzzy-Set Social Science* (Chicago: University of Chicago Press, 2000) 7–10.

⑥ 丹尼尔·卡拉曼尼著,蒋勤译:《基于布尔代数的比较法导论》,上海:格致出版社,2012年,第13页。

方法也能回答的简单问题或常识性问题。^① 多数社会科学研究依靠的是事后获得的观察值而非实验数据,因此只能通过对样本的控制来尝试达到“准实验状态”。如表 2 所示,在不考虑变量 C 的情况下,案例 1 和案例 2 形成了经典的“最大相似性”比较,由此推断变量 A 是 Y 的原因,但在现实情况多数会存在变量 C 对结果进行干扰,以至于无法用密尔方法将其排除,从而使得变量 C 成为重要的竞争性解释,只能通过理论叙述、演绎推理以及过程追踪等方式认为这些要素对于结果的影响不重要,但同时也造成了确定性的损失。

表 2 竞争性解释与求异法的或然性来源

案例	变量 A	变量 B	变量 C	因变量 Y
案例 1	1	1	1	1
案例 2	0	1	0	0

注:表 2 为笔者自制。

3. 理论的外部效度与或然性来源

理论的外部效度则是在特定范围内多案例的基础上,试图给出超越这些案例的解释力,即追求更强的“普遍性”。当理论被向外推广时,研究者则需要面临更复杂的时空情境,纳入更多的新样本。新样本中一些新变量的出现可能会导致理论结果的偏差,在此背景之下,研究者则需要权衡理论的精炼性和准确性,而精简理论的过程同样会带来确定性损失。

理论构建需在一定程度上符合基于精简性的“奥卡姆剃刀”原则。^② 因为理论是对现实的简化,以此来解释起作用的基本要素和必要的因果关系,而重要的假设应当用尽可能少的变量来解释尽可能多的现象。^③ 理论简化的过程本身是一种策略,它经常需要牺牲部分异常案例的解释力来维持理论的精简性,这也构成了理论或然性的来源。虽然减少变量来增加精简性的情况并不多见,但是研究者时常需要权衡是否“值得”通过增加变量以解释个别异常案例。如表 3 所示,在变量 C=0 的情况下,案例 1—40 是案例研究中经典的双变量模型,即变量 A 和变量 B 为 1 时,变量 Y 为 1,

其他任何情况 Y 都为 0,即变量 A 和变量 B 分别构成了 Y 的“必要条件”且两者联立时构成了 Y 的“充分条件”。假设研究者需要进一步拓展理论的外部效度,将案例 41 纳入到解释中,而深度案例研究的结果显示由于变量 C=1,导致了因变量 Y 为 0,在此情况下,研究者需要面临的选择是:要么为案例 41 单独增加一个变量从而保持 100%的解释力;要么忽视极端案例,接受理论的或然性,即解释力从 100%下降到 97.6%。显然,后者通常是更明智的选择。然而,当每种类型都只有一个案例时,那么解释力则会从 100%下降到 80%,这时加入变量 C 的必要性就有所增强。

表 3 奥卡姆剃刀、异常值与外部效度的或然性来源

案例	变量 A	变量 B	变量 C	因变量 Y
案例 1-10	1	1	0	1
案例 11-20	1	0	0	0
案例 21-30	0	1	0	0
案例 31-40	0	0	0	0
案例 41	1	1	1	0

注:表 3 为笔者自制。

4. 小结

前文从内部效度和外部效度两个方面阐述了比较案例研究的诸多步骤如何不断损失确定性而产生了或然性。如表 4 所示,或然性主要源于案例比较研究中的四个步骤:其一是对变量和事件的认知,由于多数情况下案例研究并非亲历事件,人们对实际域的认知存在或然性,在对多案例赋值时存在更多模糊性。其二是在理论构建和检验过程中,基于布尔代数的因果推断存在确定性损失,同时对于因果机制的发现和检验也仅仅是不断接近于客观实在而无法彻底接近真实域的机制。其三是在过程追踪时,证据可以不断增强对理论和假说的信念,但是主观概念无法达到绝对的充分必要性,这在存在竞争性解释或与假说不符的证据时尤为明显。其四是在界定理论适用性时,出于精简理论的目的,需要牺牲部分精确性,这也

① 赵鼎新:《社会科学研究的困境:从与自然科学的区别谈起》,《社会学评论》2015 年第 4 期。
② 参见 Hauke Riesch, “Simple or Simplistic? Scientists’ Views on Occam’s Razor,” *Theoria* 67 (2010): 75-90; 卢凌宇:《大道至简:“奥卡姆剃刀”与国际关系理论》,《欧洲研究》2018 年第 5 期。
③ Kenneth Waltz, “Evaluating Theories,” *The American Political Science Review* 91.4 (1997): 913-917; Milton Friedman, *Essays in Positive Economics* (Chicago: University of Chicago Press, 1966) 14.

是其或然性的重要来源。

表 4 案例比较研究中的内部和外部效度的统一

步骤	内部效度	外部效度
1. 变量与事件的认知	事实认定	变量赋值
2. 理论构建与检验	寻找和检验机制	基于布尔代数的因果推断
3. 因果机制与理论信念	过程证据强化信念	—
4. 适用性的进一步推广	—	理论推广及异常值处理

注:表 4 为笔者自制。

四、结语:贝叶斯逻辑与两种传承的统一

本文所要思考和澄清的核心问题是,许多定量或定性研究者对案例研究存在一定误解,即认为案例研究是基于必然性的逻辑因果而定量分析则基于概率。本文认为,无论是案例研究还是回归分析等定量研究,都是或然性的。而察看内部和外部效度的过程可以视作是一个检验假设 H_0 (特定条件 C 出现时,机制 M 对于结果没有贡献)的过程。如果最终有较大的信心可以否定 H_0 ,则可以认为机制 M 对于结果存在贡献。案例研究的目标是获得足够的证据来拒绝原假设 H_0 成立的概率 $P < 0.1$ 或 0.05 ,这在逻辑上同定量

研究的 P 检验并没有本质性的差异,区别是用过程追踪的证据而非回归分析来实现上述目标。

案例研究的或然性存在于研究设计的各个步骤之中,本文通过分析比较案例研究中或然性产生的原因,认为从事实认定与变量赋值、基于布尔代数的因果推断、寻找和检验机制、过程追踪以及理论推广和异常值处理等诸多步骤几乎都会因为确定性损失而产生或然性。这种或然性一定程度源于社会科学与自然科学的差异以及本体论和认识论之间的落差,因为社会科学中许多事件无法亲历、无法完全实现控制下的实验状态以及无法完整追踪因果机制的全过程,因而理论的主观概率都远没有达到接近必然性的程度。

因此,研究者在比较案例研究中使用密尔逻辑时,需要认识到密尔方法成立的基本前提。在社会科学中,密尔方法只能用于提升对理论的主观信心而无法产生集合论意义上的充分/必要条件。比较案例研究的或然性使得研究者应当更多地使用贝叶斯逻辑来对待小样本研究,无论是数据集观察值,还是因果过程观察值,都是可以视作一种用于调整后验概率的证据。尽管小样本研究的或然性无法被消除,但是仍然可以通过多方验证的史料、最大相似性控制、重视关键证据、寻找半负面案例以及对适用理论范围的调整等方式来减少小样本研究中的不确定性,从而提升对理论和机制的信心。

An Analysis of Probability Issues in Comparative Case Studies

Ye Chengcheng¹ Cao Hang²

(1. Institute of International Studies, Shanghai Academy of Social Sciences, Shanghai 200020;

2. Institute of the Belt and Road Studies, Yunnan University, Kunming 650091, China)

Abstract: Comparative case studies (CCS) is one of the most important issues in social science methodology. However, scholars are confused about why the Mill's method based on small samples can come to a more deterministic conclusion than the statistical analysis of large samples, and thus it is necessary to clarify the probability of CCS. Both case studies and quantitative analysis are probable, and the basic logic of the two cultures is the same. The probability of CCS stems from the differences between social sciences and natural sciences, which lead to deterministic losses in all steps of research design. The deterministic loss results from internal validity and external validity: the former includes fact finding, testing mechanism and process tracing; the latter includes variable assignment, causal inference based on Boolean algebra and the extension of the application of the theory. Consequently, CCS are more applicable to Bayesian logic, viewing causal process observations and dataset observations as evidence for adjusting subjective confidence in the theory, rather than seeking deterministic explanations on the basis of small samples.

Keywords: case study; probability; validity; deterministic loss; Bayesian logic

[责任编辑 刘 慧]