

基于历史的因果识别设计在政治学研究中的应用*

马 啸**

【摘要】 社会科学的因果识别革命使得政治学研究越来越偏重基于设计的实验方法。同时学术训练也更着重以计量为导向的因果识别、实验设计等，而轻视了对于历史知识的积累和对研究对象更深层次的了解。本文以政治制度和政治态度的有关研究为例，探讨历史事件在以识别因果关系为导向的研究设计中的独特作用。基于历史的因果识别分析对研究人员在理论、背景知识、数据分析等多个维度上都提出了更高的要求。

【关键词】 历史；因果识别；政治制度；政治态度

过去十多年里政治学实证研究经历了因果识别（causal identification）的革命。这种现象的出现，一方面是因为越来越多的经济学家开始尝试政治学议题，潜移默化地将经济学的研究范式引入了政治学，另一方面也是因为政治学的研究成果更为频繁地被作为政策决定的依据，研究人员不再满足于从数据中寻找相关性，而是希望发现因果关系。在此之前，政治学主流的研究方法是基于对观察数据（observational data）的分析。无论是通过个案比较，还是运用大样本统计分析，本质上都是通过案例间的比较排除干扰因素，以发现因变量与自变量之间的相关性。但是基于观察数据的因果关系分析具有较大的局限性。即使是非常稳健的相关性的结果，也可能会受到包括样本选择偏差、反向因果、遗漏变量等方面的质疑。

这种变化对政治学的研究范式和研究问题都产生了深远的影响。为了应对观察数据在因果关系识别上的挑战，政治学研究日益青睐基于设计的实验法。研究人员通过严谨的实验控制对照组设计，人为地改变实验组自变量的值，并比较其与对照组因变量的区别。因其研究设计的严谨性，实验法可以为研究者提供非常精准的因果效应的估计，其结果也往往更具有政策参考的价值。但是实验法同样具有局限

* 本文得到教育部人文社会科学重点研究基地重大项目“国家治理经验评估”（编号：16JJD810001）和教育部中央高校基本科研业务费专项资金资助。

** 马啸，北京大学中国政治学研究中心助理教授。

性。首先并不是所有的研究议题都可以用实验法来检验，而且往往那些无法用实验法检验的话题，恰好具有重要的理论意义。例如，如果想要研究民主化对一国经济发展的影响，研究人员显然无法将威权国家随机分组，让一部分国家成为民主国家并观察其经济发展表现。有些问题虽然理论上可以通过实验的方法进行检验，但是研究设计受道德考量的制约也无法实现。例如，如果想要知道战乱经历对人的政治态度的影响，研究人员就需要让受访者去战场上亲身体验一下战争的残酷，并测量政治态度的变化。但这种可能伤害到受访者的研究设计显然无法在现实中执行。这些局限使得使用实验法的政治学研究者逐渐回避社会科学领域的传统经典理论问题（例如现代化理论、国家构建、民主化、冲突等），转而进行一些以政策为导向的微观研究，研究的话题和所涉及的理论文献向应用公共管理靠拢。长期看来，强调因果识别的研究导向和实验法的使用，使得政治学的研究出现以数据和研究方法为驱动（data and methods driven）的倾向，并且减少了与传统经典话题的真正有意义的对话。

本文认为，政治学的一些传统研究方法在以因果识别为研究导向的潮流中可以焕发新的活力。这里说的传统研究方法，特指比较政治学中以国别、区域为研究对象的研究，要求研究者对研究对象的历史进行深度了解和发掘，对所研究对象的社会背景知识（不仅限于自己所研究的特定话题）有广泛积累。本文认为，掌握研究对象的历史与细节知识，本身并不与因果识别相矛盾。相反，掌握丰富的关于研究对象的质性知识，为研究者通过“自然实验”（natural experiment）来识别因果关系打开了大门。与基于设计的实验相比，自然实验具有成本低廉、可操作性强、可以检验的理论话题更为广泛等优势。这种方法将历史事件所造成的长期冲击纳入因果识别的框架，来分析制度、观念等内生性较强，但同时又颇具理论意义的变量的影响，是一种综合了质性与定量方法、观察数据与实验设计优点的研究路径。同时，基于历史事件的自然实验法对学者提出了更高的学术训练要求。研究者不仅需要深入了解研究对象的背景历史知识，也需要熟练掌握计量分析技巧，并将研究设计与重要的理论问题相结合。可以说，这种研究方法对选题立意、证据收集以及数据分析这几个评价研究的重要维度均有强调，同时不以牺牲其中一方来成就另一方。当然任何研究方法都不是完美的，基于历史事件的因果识别同样具有其局限性。

本文首先简单介绍历史自然实验的概念，之后分别结合制度变迁和政治态度两类具体研究，对基于历史事件的因果识别设计进行阐释，最后对这种方法的局限性做简单的讨论。

一 历史自然实验与因果识别

使用观察数据并通过相关性分析得出结论的研究面临着两大挑战。第一个挑战是反向因果问题，即在分析中被认为是解释变量的因素也可能受到被解释变量的影

响而发生改变。第二个挑战是遗漏变量的干扰，即解释变量和被解释变量可能本身不存在因果关系，而是受某个未观测到的变量影响而同时发生改变。^① 研究者尝试使用各种方法解决这些问题，包括在分析中使用解释变量的时间滞后项以及在回归中控制尽可能多的干扰因素。但是这种“事后补救”（*ex post*）的办法存在各种局限，无法真正解决问题。^②

为了解决这些问题，计量经济学家提出了工具变量的方法。^③ 简而言之，工具变量是一个与被解释变量不直接相关，而通过影响解释变量来改变被解释项的变量。工具变量需要满足排他性限制（*exclusion restriction*），即其只能通过改变解释变量这一个渠道，而非其他途径，来影响被解释变量。为了满足这一条件，研究者采用的工具变量大都基于偶发的事件、自然灾害，或者政策的人为断点等。其背后的逻辑与基于设计的实验法相近，即通过近似随机（*as-if random*）的干预，排除其他干扰因素对被解释变量的影响。

历史事件形成的冲击同样可以被用于构建工具变量。虽然事故或自然灾害造成的短期冲击也经常被用于因果识别，但其检验的对象往往较为简单，容易受外界影响发生变化，且易于测量（例如选举结果等）。^④ 而很多研究者感兴趣的对象，例如制度等，则更为复杂且不易变化，所以使用此类外生冲击往往具有局限性。而重大的历史事件，则可以在更长的时间维度上积淀出更深远持久的影响，因此可以用来检验一些更为复杂且不容易改变的事物的因果效应。下面就以制度和政治态度的研究为例，探讨历史事件作为工具变量在因果识别设计中的运用。

二 基于历史的因果识别在制度研究中的应用

制度变迁的原因及其影响是政治学关注的一个核心话题。自从 20 世纪 50 年代现代化理论提出后，主流观点认为政治制度的发展是经济发展的结果。^⑤ 20 世纪 70 年代以后，随着新制度主义学派的兴起，人们对制度与经济发展间的因果关系又提出了新的问题。制度主义学派认为良好的制度是经济发展的前提而非结果，这与现代化理论的因果关系形成了直接的冲突。由于数据的局限，尽管经济学家和政治学

① 另一种导致分析偏误的常见原因是选择偏差问题，也可以被归为遗漏变量问题。

② 关于“事后补救”（*ex post*）型和“事先设计”（*design-based*）型因果识别的区分，见 Thad Dunning, *Natural Experiments in the Social Sciences: A Design-based Approach*, Cambridge University Press, 2012。

③ Joshua D. Angrist and Alan B. Krueger, “Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments,” *Journal of Economic Perspectives*, 2001 (15), pp. 69 – 85.

④ 运用事故和自然灾害造成的短期影响的研究，见 Christopher H. Achen and Larry M. Bartels, “Blind Retrospection: Electoral Responses to Drought, Flu, and Shark Attacks,” *Working Paper*, 2004。

⑤ Seymour Martin Lipset, “Some Social Requisites of Democracy: Economic Development and Political Legitimacy,” *American Political Science Review*, 1959 (53), pp. 69 – 105.

家尝试采用各种高级的计量方法，但还是没有办法取得共识。^① 分析的难点集中在两处：第一，政治制度的决定因素可能是多种多样的（包括历史、文化、气候、资源、战争、宗教等），利用观察数据只能控制有限的干扰变量，无法穷尽所有可能影响制度的因素；第二，经济发展与制度变迁这两个核心变量在理论上可以互为因果，即经济发展既可以导致制度变迁，同时制度变迁也会对经济发展产生影响。因此即使在回归分析中发现两者具有非常稳健的相关性，也无法判定经济发展与制度变迁在现实中的因果关系。

阿西莫格鲁等人于2000年发表的论文《比较发展的殖民地起源》，将历史遗产（即各国的早期殖民模式）用作识别制度变迁的外生冲击，是一项具有划时代意义的重要研究。^② 阿西莫格鲁等人发现，曾经被欧洲殖民者统治的国家间存在巨大的经济发展水平差异和制度差异。而这些国家间的制度差异可以用第一批殖民者抵达殖民地时的死亡率来解释。在那些死亡率较低的区域，殖民者会选择长期驻扎，将殖民地当作自己的家园并引入欧洲的法律和产权保护等制度。而在那些早期殖民者死亡率较高的区域，因为环境恶劣、不适合长期居住，殖民者将殖民地变成了榨取资源和劳动力的场所。现代国家制度的确立和发展受到了这种早期殖民地差异的路径依赖影响。那些早期殖民者死亡率较低、殖民者大规模定居的国家，逐渐发展出了一套较为公平、完善的法治体系，这是现代民主国家的雏形。而那些殖民者死亡率较高、殖民者没有大规模定居的地方，殖民地时期的榨取性制度安排则演变为现代专制政权的雏形。早期的殖民者死亡率不可能受现代发展水平的影响，也不太可能通过后续政治制度演变这一渠道以外的机制来影响现代经济发展水平，因此其符合作为制度工具变量的排他性限制。这是新制度主义的核心观点，即完善的产权保护制度是经济发展的前提，这项研究提供了系统性的、跨国的因果识别检验，因此具有重要意义。

但是阿西莫格鲁等人的结论不是一蹴而就的偶然发现，而是建立在历史学家和区域研究专家大量的史料发掘基础之上。其分析中用到的80多个原欧洲殖民地国家早期殖民者死亡率的数据，就是基于历史学家菲利普·柯廷（Philip Curtin）关于欧洲殖民史的30多年的研究成果。^③ 而早期殖民者的定居模式与后续制度发展之间的关系，也非阿西莫格鲁等人的首创。经济史学家索克罗夫和安格曼（Sokoloff

① 例如：Robert E. Hall and Charles I. Jones, "Why do Some Countries Produce So Much More Output Per Worker than Others?" *The Quarterly Journal of Economics*, 1999 (114), pp. 83 - 116; Adam Przeworski, Michael E. Alvarez, Jose Antonio Cheibub, and Fernando Limongi, *Democracy and Development: Political Institutions and Well-being in the World, 1950 - 1990*, Cambridge University Press, 2000.

② Daron Acemoglu, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 2001 (91), pp. 1369 - 1401.

③ Philip D. Curtin, *The Atlantic Slave Trade: A Census*, University of Wisconsin Press, 1972; Philip D. Curtin, *Death by Migration: Europe's Encounter with the Tropical World in the Nineteenth Century*, Cambridge University Press, 1989.

and Engerman) 在研究南北美洲巨大的发展差异时, 就注意到了两地不同的环境和自然禀赋 (factor endowment) 对欧洲殖民者建立早期殖民地政权的影响。^① 南美洲炎热湿润的环境使得经济作物的产量较高。而经济作物对劳动力需求巨大, 因此殖民者在当地建立的初始制度就是以榨取为目的, 通过奴役当地人口实现对自然资源和劳动力的双重剥削。相比之下, 北美洲温和的气候环境适合种植对劳动力需求并不高的小麦等作物, 因此并没有大范围地出现像南美洲那样的以奴役人口为目的的榨取性制度安排。而这种初始的殖民地制度对两地后续的制度发展产生了深远的影响。阿西莫格鲁及其合作者的贡献就在于巧妙地将现成的史料与计量方法结合, 运用工具变量的研究设计检验了制度对经济发展的因果效应。

对阿西莫格鲁等人研究的众多批评之一是其结论的外部效度 (external validity)。西方国家的殖民历史是一次制度迁移的宏大自然实验, 殖民政权的建立对于殖民地的政治制度构成剧烈的外生冲击, 很多殖民地在此之前甚至不存在现代意义的国家。而那些没有彻底沦为殖民地的区域 (例如东亚和中东地区等), 则没有经历类似的外部引入制度的冲击, 因此也没有进入阿西莫格鲁等人的分析样本。对于这些历史悠久, 而且政权能力足以抵御西方殖民者、不至于彻底沦为殖民地的国家, 如何评价制度与经济发展间的关系值得思考, 他们的制度设计是否也如新制度主义学派所认为的那样, 对当地的经济起到了关键的作用?

学界为此已经进行了不少有意义的尝试。梅拉尔多 (Menaldo) 在研究中东与北非国家时, 发现那些在“第二次世界大战”以后仍然保持了君主制的国家 (例如卡塔尔、阿联酋、沙特阿拉伯等) 普遍政治稳定、经济发展水平较高。而那些名义上实行了共和制的国家 (如伊拉克、叙利亚、利比亚等), 则被政治动荡所困, 经济发展水平不高。^② 一般认为共和制相比君主制是一种更具有现代性的制度, 采用了共和制的国家在治理的效果上也应该更好, 为何在中东地区出现了相反的情况? 梅拉尔多经过细致的案例研究后发现, 那些政体形式上采取了共和制的中东国家在本质上与君主制国家一样, 都属于专制政权。而不同的是, 君主制国家在权力继承、资源分配、政策制定等重要议题上, 具有更为明确的行事规则 (如长子继承制、王室会议等), 而且君主制也通过将传统的神秘化保护这些规则不被轻易修改。相比之下, 那些采用了共和体制的专制国家, 在上述重大议题上缺乏规则, 即使形成了规则也容易被强人领导所改变。而缺乏明确的规则使得这些国家的统治精英更容易出现分裂, 进而影响这些国家的政治稳定与经济发展。通过收集战后中东国家发展的面板数据, 梅拉尔多发现实行了君主制的国家的确在经济发展的各项指标上比共和制国家表现更好。

① Kenneth L. Sokoloff and Stanley L. Engerman, “Institutions, Factor Endowments, and Paths of Development in the New World,” *Journal of Economic Perspectives*, 2000 (14), pp. 217–232.

② Victor Menaldo, “The Middle East and North Africa’s Resilient Monarchs,” *The Journal of Politics*, 2012 (73), pp. 707–722.

梅拉尔多的分析中面临的一个难题是如何证明君主制与较高经济发展水平间的相关性是因果关系。人们对这一反直觉的实证关系可能会进行各种质疑。较高的经济发展水平也有可能是中东君主制国家能够成功抵制要求实行共和制诉求的一个重要先决条件。此外,也可能存在一个尚未被了解的变量(例如,一国的资源丰富程度),既决定了一个国家的经济发展水平,又影响该国是否能够维持君主制的统治。为了解决这些问题,梅拉尔多尝试寻找能够解释中东国家政体形态的工具变量。通过对历史和人类学的广泛研究,他发现在该区域实行君主制的国家,历史上长期的社会形态以游牧部落为主。而君主政权早期的贵族则是由各个游牧部落的主要头领构成。在那些实行共和制的国家,历史上则长期以城邦为主,游牧部落头领对政治的影响较小。通过阅读农业史和气象史的资料,他进一步发现,气候和土壤条件是决定一地社会形态的主要原因。在那些降水充沛、土壤肥沃的区域(如伊拉克的两河流域),从很早开始就具有发达的农业,能够承载更多的人口,因此也更早地出现了城市。而在降水稀少、土壤贫瘠的地区,游牧则成为主要的生产生活方式。通过收集中东地区历史降水量的数据,梅拉尔多找到了影响该地区当前政体形态的工具变量,证实了特定的政体形态(君主制)对威权国家的经济发展有重要的影响。^①事实上,将气候等自然因素作为制度起源的研究,在社会科学内有着悠久的传统,^②但直到最近才被系统性地纳入因果识别的框架。

上述两项研究都是基于跨国的制度比较,而类似历史方法在单个国家内的跨区域分析(sub-national analysis)中同样应用广泛。比较而言,在一国内的分析,因为控制了诸如传统、文化、语言等潜在的干扰因素,研究人员能够更好地分离和识别制度变化所带来的影响。^③例如,杨(Yang)关于科举制度的长期影响的研究发现,那些在明代考取进士越多的地区,在现代人均收入也越高、基础设施越完善、对教育的投入也越多。^④其认为这一相关性的主要因果机制是大量进士产生后形成的地方尊重人力资本投资的传统,而对教育的重视直接导致这些地方在现代发展水平更高。为了识别历史进士人数对当代发展水平的因果效应,杨通过发掘县志档案,将明代各地自然灾害(异常降水)的记载用作当地考上进士人数的工具变量。在另一项研究中,章奇和刘明兴发现浙江各地私营经济发展水平与该地在改革开放初期本地干部占比呈现正相关。他们认为本地干部在政治上相比空降的外来干部更

① Victor Menaldo, *The Institutions Curse: Natural Resources, Politics, and Development*, Cambridge University Press, 2016.

② 例如20世纪50年代魏特夫关于水利和东方专制主义的研究,见Karl August Wittfogel, *Oriental Despotism: A Comparative Study of Total Power*, Yale University Press, 1959.

③ Richard Snyder, "Scaling Down: The Subnational Comparative Method," *Studies in Comparative International Development*, 2001 (36), pp. 93-110.

④ Clair Zhuqing Yang, "Long-Run Persistence and Interrupted Development: Evidence from Historical China," *University of Washington Working Paper*, 2017.

为弱势，因此需要通过保护本地的私营经济来获得基层政治支持。^①章奇和刘明兴用浙江省各县1949年以前本地游击队的分布数据作为中华人民共和国成立后本地干部比重的工具变量，证实本地干部比例与私营经济发展具有因果效应。受该项研究启发，本文作者尝试解释浙北、苏南两地在改革开放以后经济发展模式差异的历史决定因素。本文作者发现，苏南地区之所以高度依赖外资，与改革开放初期该地国有、集体经济发达，而当时国家法律又要求外商投资必须与本地企业合资有关。国有与集体企业匮乏的浙江不具备吸引大量外资的条件，而浙江的地方官员在考核的压力之下最终只能通过发展本地的私营经济来完成考核目标。通过对地方历史的进一步发掘，本文作者发现两地在改革开放初期国有、集体经济的发达程度与两地的地理位置有关。浙江因地处对台海防前线，在计划经济时代接受的国有投资远远少于江苏。因此笔者将两地受台湾军事威胁的程度（至台北距离）作为工具变量，发现两地改革开放初期国有资本的充盈程度的确和两地如今的发展差异间存在着因果效应。^②

这种方法在中国之外也有广泛的运用。例如，梅丽莎·戴尔（Melissa Dell）等人的研究发现，越南历史上受中国影响较大的北部地区，基层治理中多采用更为制度化的村庄结构。与之形成对比的是越南南部地区，基层治理更多依靠非正式的人际网络和恩庇侍从关系。这种历史制度差异对当今越南的区域发展差异产生了重要影响。北部地区的基层社会在原有的制度框架的基础上，能够更好地克服集体行动的困境，通过公民自治来解决公共产品提供和再分配的问题。而南部地区则不具备类似的制度基础，发展也相对较为迟缓。通过运用断点回归的研究设计，戴尔等人发现基层制度的历史遗产对越南当今的区域发展有显著的影响。就如同上述例子所展现的那样，基于一国内的因果识别设计往往较跨国分析研究设计更为精致，测量更为准确，更容易获得可靠的结果，但同时对研究者也提出了更高的要求。研究者除了掌握基本的计量分析技巧之外，还需要对所研究国的历史、文化、语言、史料的可得性等都有相当深入的了解。

三 基于历史的因果识别在政治态度研究中的应用

基于历史的因果识别方法另一广泛使用的领域是对政治态度的研究。尽管政治制度与政治态度的研究经常被认为是两个壁垒分明的学术流派，但广义的政治制度也包括了能够引导人们行为的一些非正式制度，包括观念、文化、信任等态

① 章奇、刘明兴：《权力结构、政治激励和经济增长——基于浙江民营经济发展经验的政治经济学分析》，格致出版社，2017。

② Xiao Ma, "Befriending the State? Socialist Legacies, Local State Sector, and Foreign Direct Investment in China," A Paper Presented at Fudan University, Shanghai, China, May 18, 2018.

度元素。^① 因此从作为研究对象的角度看,两者也并非全无共通之处。在识别因果机制这一维度上,两者面临着一些共同的挑战。首先,两者都属于稳定的均衡,因此不容易被轻易改变,具有一定的延续性。^② 这增加了研究者寻找能影响政治态度的外生冲击的难度。其次,政治态度本身的决定因素也较为多样,包括个体的年龄、受教育水平、家庭环境、历史经历等。这些因素不仅互相之间存在相关性,而且经常和其他的被认为能够决定政治态度的因素相关。这给准确地识别政治态度的决定因素和其对其他行为的影响制造了不少挑战。研究者利用具有重大影响的历史事件来识别政治态度的决定因素。这些事件的发生(及其终止)不受被影响个体的控制,因此其对个体的影响可以看作是近乎随机的干预。而且这些事件往往存续时间较长,其干预的效果足够改变个人及群体的观念态度的均衡值,形成深远持久的影响。

纳恩(Nunn)和沃彻肯(Wantchekon)在一项研究中发现,欧洲殖民非洲期间受奴隶贸易影响较大的族群个体,在2005年的非洲价值观调查(Afrobarometer survey)中普遍表现出对他人(包括亲属、邻居、同族、地方政府等)较低的信任。^③ 尽管这一相关性发现符合心理学、文化人类学等对低信任水平的已有解释,但是也面临着因果识别的挑战。这一相关性也有可能因为当时欧洲殖民者策略性地选择了那些信任程度一直较低的非洲族群进行奴隶贸易,而并非是已经结束了上百年的奴隶贸易对当前的信任水平造成了负面影响。为了识别奴隶贸易对政治态度的影响,纳恩和沃彻肯使用非洲各地到海岸线的距离作为殖民时期奴隶贸易强度的工具变量,并发现奴隶贸易的强度对个体的信任态度有显著的负效应。他们同时发现,各地距离海岸线的直线距离仅在非洲与政治信任存在负相关性,而在欧洲和亚洲则不存在类似相关性,说明与离海岸线的距离只有通过历史上的奴隶贸易而非其他现当代的渠道影响政治信任,为工具变量满足排他性限制提供了有力证据。

基于历史的因果分析同样可以为反直觉的发现提供有力的佐证。科恩和海恩穆尔勒尔(Kern and Hainmueller)在一项研究中发现,在原民主德国(东德)的居民中,有观看西德电视节目经历的居民对东德政权的支持度更高,对生活现状的满意度也更高。^④ 这与文献中关于外国媒体对威权国家政治态度的影响的已有研究结论

① Douglass C. North, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, 1990; Avner Greif and David D. Laitin, "A Theory of Endogenous Institutional Change," *American Political Science Review*, 2004 (98), pp. 633 - 652.

② 当然在美国政治的研究中,也有一种观点认为政治态度可以受外界影响而发生短期的改变。例如一些事故新闻等会对人们的投票行为产生影响,参见 Christopher H. Achen and Larry M. Bartels, *Democracy for Realists: Why Elections Do Not Produce Responsive Government*, Princeton University Press, 2017。

③ Nathan Nunn and Leonard Wantchekon, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011 (101), pp. 3221 - 3252.

④ Holger Lutz Kern and Jens Hainmueller, "Opium for the Masses: How Foreign Media Can Stabilize Authoritarian Regimes," *Political Analysis*, 2009 (17), pp. 377 - 399.

相反。科恩和海恩穆尔勒尔认为形成这种结果的主要机制是西方媒体为东德居民提供了娱乐消遣（因此提高了满意度），而并非其他遗漏变量驱使的结果。他们利用两德分立时期西德电视信号发射塔的位置及信号覆盖范围作为识别手段，发现分布在信号覆盖范围两侧的居民的政治态度存在显著的差异，为上述发现提供了有力的支持。

中国研究政治态度的学者同样将目光投向了历史。因为中国近代史的多次动荡和历史数据的相对缺乏，绝大部分研究者将可能影响人们政治态度的历史事件范围缩小到了中华人民共和国成立以后。例如，陈济冬等发现，在土地改革时期，家庭成分被定义为“坏阶级”（例如“地主”等）的人士及其后代，当前对政府的信任度普遍较低，尽管这些人在改革开放后取得了较高的社会经济地位。^①他们进一步用土壤适耕性等指标作为被定为“坏阶级”的工具变量，发现上述结果稳健。与之类似的是，王裕华的研究也发现，20世纪60年代后期武斗中伤亡越严重的区域，居民当前对政府的信任程度越低。^②为了检验武斗对政府信任的因果效应，王裕华使用了硫矿所在位置作为武斗激烈程度的工具变量。军工厂多靠近硫矿所在地，有解放军驻守并维持秩序，因此出现大规模武斗伤亡的概率较低。这些研究所共有的假设是，政治态度一般情况下较为稳定，除非经历剧烈外部冲击（如社会动荡、阶级斗争等），否则很难形成新的均衡。而新的均衡一旦形成，又延续了先前的稳定性并能够在代际传承。这种稳定的特点与制度具有相似性，因此在作为研究对象的方法上也具有共性。

四 结论与讨论

本文结合若干研究实例，着重介绍了历史事件在社会科学研究因果推断中的作用。这种研究范式为处在因果识别革命进程中的政治学研究提供了一种新的路径。研究者无需在定性和定量、理论和方法中做出非此即彼的选择。基于历史的因果分析要求研究者同时具有对研究对象质性细节（特别是历史背景文化）的关怀和对严谨研究设计与定量研究方法的掌握。这种研究范式将原本被研究者所逐渐抛弃的质性信息重新纳入严谨的分析框架，扩展了强调因果识别设计的研究可以检验的理论广度。从某种意义上，找到一个可供识别因果关系的历史事件犹如找到一件被历史尘封已久的宝物，是社会科学研究中充满了乐趣的一种经验。

任何一种研究方法都存在其局限性，基于历史的因果识别也不例外。首先，将历史事件的外生冲击作为工具变量时，较难证实其排他性限制。在漫长的历史进程

① Jidong Chen and Anna Zhang, "Embittered Winners: Right Deprivation and Ideological Distortion," *Working Paper*, 2018.

② Yuhua Wang, "For Whom the Bell Tolls: The Political Legacy of China's Cultural Revolution," *Harvard University Working Paper*, 2018.

中，一个事件的后续发展可能通过各种渠道对当前产生影响，如何证明该事件要通过研究人员所关心的特定因果机制对当前产生影响是一大挑战。因为历史数据的稀缺性，很多时候研究人员只能依赖文字描述（而非数据）来回应对排他性限制的质疑，不能给出非常令人信服的辩护。同时，因为历史数据的稀缺性，研究人员经常缺乏能检验同一事件的交叉证据，因而数据中存在的衡量偏差往往无法得到检验和修正。而这些衡量偏差有时会对分析结果产生至关重要的影响。^①其次，这种研究方法对于学术训练也提出了更高的要求。发现一个好的历史自然实验，除了掌握计量方法外，还需要研究人员在历史、政治、地理、经济领域有广泛而且深厚的积累，而且同时具有将历史实证数据与理论话题相互联系的直觉（intuition）。而这种直觉，很难通过短期高强度训练获得，是研究人员长期积累的结果，甚至有可能是一种学术界培训之外的能力。好的历史自然实验的不可得性，导致部分研究出现了新的数据驱使的趋势，即研究话题受可以获取数据的历史冲击的限制，没有形成连贯的研究议题。如何在历史数据和理论议题之间建立起有效的对话，是使用这种方法的研究人员所要面临的一大挑战。

History-based Causal Inferences in Political Science Research

Ma Xiao

Abstract: The causal identification revolution in the social sciences has led political scientists to conduct more design-based experiments. Research training in political science now puts more emphasis on causal inferences and experimental designs, neglecting the accumulation of historical and contextual knowledge about the research subject. This article, based on studies of political institutions and attitudes, argues that historical knowledge plays an important role in the design of causal inference research. Such a method requires that political scientists have more balanced expertise in theory, contextual knowledge, and analytical methods.

Keywords: History; Causal Inferences; Political Institutions; Political Attitudes

^① 例如，关于阿西莫格鲁等的殖民地制度起源的研究，就有学者提出疑问，认为其数据衡量存在偏差。参见 David Y. Albouy, “The Colonial Origins of Comparative Development: An Empirical Investigation: Comment,” *American Economic Review*, 2012 (102), pp. 3059–3076.