

DOI: 10.16538/j.cnki.fem.20191105.001

因果关系的实地实验与新实证发展经济学 的贫困治理之道

——2019年度诺贝尔经济学奖得主主要经济理论贡献述评

李宝良¹, 郭其友²

(1. 华侨大学 经济与金融学院, 福建 泉州 362021; 2. 厦门大学 经济学院, 福建 厦门 361005)

摘要: 阿比吉特·班纳吉、埃丝特·迪弗洛和迈克尔·克雷默三位经济学家因为对减轻全球贫困的实验方法所做的突出贡献, 被瑞典皇家科学院授予2019年度诺贝尔经济学奖。本文从实地实验方法和贫困问题的具体研究两个方面对他们的突出贡献进行述评, 得到四个主要的结论: 第一, 在对比自然科学和社会科学关于因果关系实证研究之异同的基础上, 阐明实地实验与自然科学和其他社会科学有着相同的基本原理, 并且都面临内部有效性和外部有效性的争议。第二, 实地实验作为因果关系推断的工具在经济学实证研究中有着特别的优势, 应该将其纳入经济学家因果关系分析的工具箱中; 此外, 实地实验可以作为政府制定经济政策和企业经营管理实践的有效工具, 应该予以推广。第三, 他们对贫困问题所进行的实地实验以及由此得到的结论具有重要的参考借鉴价值, 这一点值得肯定。特别是, 在机制设计或者政策制定的过程中, 细节的设计非常重要, 必须仔细加以考虑。第四, 从以亚当·斯密为代表的古典经济学家的经济发展思想到现代发展经济学的演化历程及中国扶贫的实践来看, 他们将贫困问题作为发展经济学研究的核心使得发展经济学研究的视野过于狭窄, 发展经济学应该回归和重视经济发展机制的研究。

关键词: 阿比吉特·班纳吉; 埃丝特·迪弗洛; 迈克尔·克雷默; 实地实验; 因果关系; 贫困与扶贫; 发展经济学

中图分类号: F270 **文献标识码:** A **文章编号:** 1001-4950(2019)11-0136-17

一、引言

贫困问题是困扰全球特别是发展中国家的最为棘手的问题之一, 也是发展经济学研究的核心问题之一。然而, 发展经济学家对贫困的成因及其对策持有两种截然不同的观点。一种是

收稿日期: 2019-11-05

基金项目: 福建省社会科学规划一般项目(FJ2018B057)

作者简介: 李宝良(1980—), 男, 华侨大学经济与金融学院副教授, 经济学博士(通讯作者);

郭其友(1963—), 男, 厦门大学经济学院教授, 经济学博士。

以贫困问题专家杰弗里·萨克斯为代表的发展经济学家,他们承袭了庇古和凯恩斯的政府干预的思想,认为穷国之所以穷,是因为陷入所谓的“贫困陷阱的恶性循环”,单凭这些国家的力量难以从这种恶性循环中挣脱出来,因此,国际社会应该伸出“扶持之手”,以援助等方式帮其跳出贫困陷阱的泥潭(杰弗里·萨克斯,2007);另一种以美国经济学家威廉·伊斯特利为代表,他们继承亚当·斯密“看不见的手”的衣钵,认为穷国之所以贫穷,是由于缺乏正确的激励机制,只要把激励搞对,自然会有一只“看不见的手”会将穷国带出贫困陷阱,事实上,持有这种观点的发展经济学家根本不相信贫困陷阱的存在(威廉·伊斯特利,2005)。^①

基于贫困的这两种不同观点,发展经济学家对许多解决贫困问题的政策建议也有截然相反的看法。以外国援助为例,持“扶持之手”的贫困专家实际上是把贫困问题视为市场失灵的必然结果。因此他们认为,如果国际社会不施以援手,这些国家就难以从贫困陷阱中解脱出来,就像政府不通过补贴和税收等方式就难以纠正市场失灵一样。要将贫困的人解救出来需要更多的资源,而发展中国家往往缺乏必要的资源,因而,他们认为外国援助多多益善,涵盖了粮食、健康、教育、基础设施等各个领域。与之相对的是,持“看不见的手”的发展经济学家则反对任何形式的外国援助,他们认为如果不把激励机制搞对,再多的外国援助也无济于事;很多发展经济学家甚至认为,如果不把激励机制搞对,更多的外国援助反而会起逆向激励的作用,对贫困问题的解决有害无益。

这两种针锋相对的观点都失之于笼统。持“扶持之手”的发展经济学家认识到了外国援助的作用,但缺乏对援助穷人的有效手段和方式的具体研究,使得外国援助常常被浪费而无法起到其预期的目标,甚至与最初的目标背道而驰;而持“看不见的手”的发展经济学家则认识到了激励机制的重要作用,然而由于缺乏对激励机制如何起作用缺乏具体的了解,使得他们对如何才能将激励机制搞对也未能提出切实可行的措施建议。事实上,使援助有效地发挥作用归根结底是把激励机制搞对的问题。而要把激励机制搞对,关键在于了解人们这样或那样行事背后的动机。激励机制是一种很微妙的东西,它发挥作用的方式有时候与经济学理论的预期不一样。如果不搞清楚人们这样或那样行事背后的动机,那么设计出来的激励机制可能无法达到预期的目的,甚至适得其反。

贫困问题解决机制设计或者政策制定的关键正在于此。只有搞清楚处在贫困中的人们为什么这样或那样行事,才能设计出有效的激励机制,使得扶贫政策切实有效地发挥作用。为此,20世纪90年代中期以来,以班纳吉、迪弗洛和克雷默为代表的新一代发展经济学家将实验方法引入发展经济学研究之中,以此来探讨有效的贫困问题治理之道。这种实验研究方法以往发展经济学的研究方法不同,它不是将贫困作为一个宏观问题借助数学方法进行建模分析,而是将贫困分解为一系列具体的小问题,引入信息经济学、激励机制设计理论以及行为经济学等微观经济分析方法对这些小问题进行分析,并设计出各种可能改善或解决这些影响贫困的小问题的机制和方案,然后通过精心设计的实地实验对这些机制或方案的效果进行细致的研究,从而试图找出治理贫困问题的最有效的施策方案。这种实地实验的方法的本质是一种因果关系检验的方法,旨在找出贫困治理的机制及其结果之间的可靠的因果关系。在他们的这种方法推动下,发展经济学进入一个新的发展阶段,被称为新实证发展经济学^②(巴苏,2007)。瑞典皇家科学院在其新闻公报中称,他们开展的研究极大地提高了我们与全球贫困问题进行斗争的能力,在仅仅二十多年的时间里,他们基于实验的新方法使得发展经济学面貌一新,并且成为一

^①限于篇幅,本文未详尽地追溯双方争议的来龙去脉。事实上,“贫困陷阱的恶性循环”理论可以追溯到20世纪50年代,1953年纳克斯提出了“贫困恶性循环”理论,1956年纳尔逊提出了“低水平均衡陷阱”理论,1957年缪尔达尔提出了“循环积累因果关系”理论,从不同的角度论证贫困陷阱产生的根源。此外,反对政府干预以解决贫困问题的经济学家也不在少数,这些经济学家相信市场之能,例如芝加哥学派代表人物,1992年诺贝尔经济学奖得主加里·贝克尔就主张“给穷人创造自由市场是解决贫困的良药”。在此,感谢匿名审稿专家的中肯意见。

^②巴苏使用新实证发展经济学来指代发展经济学的一种趋势,即利用可控的随机实验或精心挑选的工具变量来揭示经济变量之间的因果关系。巴苏对发展经济学及其问题的分析可参考《比较》(第28辑),吴敬琏主编,中信出版社,2007年1月。

个欣欣向荣的研究领域。

他们的实地实验方法对经济学的实证研究和经济实践有重要的意义,他们对贫困问题进行实验所得的结论对扶贫工作有重要的参考借鉴价值。本文将从两个方面对他们在减轻全球贫困的实验方法方面所做的突出贡献进行述评。首先,在对比自然科学与社会科学对因果关系问题研究异同的基础上,结合中国经济改革和企业经营管理实践,论证实地实验在因果关系实证研究中的地位及其实践价值。其次,将贫困问题的研究放在以亚当·斯密为代表的古典经济学家的经济发展思想到现代发展经济学的演化脉络中,在结合中国扶贫实践的基础上,对他们的贫困问题实地实验的得失进行评价,进而指出发展经济学研究应该注重的方向。

归纳起来,本文得到四个主要结论。第一,不管是自然科学还是社会科学,因果关系的实证研究都有相同的基本原理,而且对因果关系的统计推断都存在内部有效性和外部有效性两方面的争议。自然科学通常使用实验对因果关系进行检验。社会科学中的因果关系实证研究也有相同的问题,然而由于学科性质的不同,具体的做法存在很大差异。经典的计量经济学方法使用多元线性回归模型或者借助偶然实验对因果关系的内部有效性进行推断,并通过稳健性检验证明结论的外部有效性;而实地实验则直接按照自然科学的规范进行实验来对因果关系内部有效性进行统计推断,实地实验的结论要加以推广,同样要考虑其外部有效性的问题。

第二,实地实验方法和其他的计量经济学分析工具一样,都是因果系统统计推断的一个工具,然而实地实验具有其他工具所不具备的优势,可以使得经济学家对经济问题背后的因果关系进行主动的设定,从而有效地解决传统多元线性回归模型中的内生性问题,同时使得经济学家不再受制于偶然实验的可得性,从而开辟了经济学因果关系检验的新空间。经济学家应该重新审视实地实验方法,将其纳入因果关系检验的工具箱中,因其具有不同于其他工具的功效。此外,我国的经济改革一直以来都秉持“摸着石头过河”的哲学,通常是通过各种政策试点来对其实施的各种政策的有效性及其问题进行分析。在这个意义上,我国的经济改革可以说是一场史无前例的大规模实地实验;企业经营管理也是如此,企业管理者经常要对经营策略、产品和服务创新的方案的有效性进行测试,这也是一种实地实验检验。因此,实地实验可以作为政府经济政策制定和企业经营管理实践的有效工具加以推广。

第三,他们对贫困问题所进行的实地实验以及由此得到的结论,具有重要的参考借鉴价值,这一点值得肯定。他们对贫困问题的研究一方面扩展了微观经济理论在发展经济学中的应用,加深了我们对经济理论本身的有效性的认识;另一方面也使得我们加深了对贫困人口的动机的了解,使得我们对某些贫困问题的解决有切实有效的具体建议。更重要的是,这些研究表明,要使所设计的激励机制或政策有效,细节的设计非常重要:谁、做什么、在什么时候做、在哪里做、为什么要这么做以及我们可以提供什么样的激励,所有这些问题在进行机制设计时都必须仔细加以考虑(尤里·格尼茨和约翰·李斯特,2015)。这些对我国扶贫工作的进一步实施都有直接的参考借鉴价值。

第四,从以亚当·斯密为代表的古典经济学家的经济发展思想到现代发展经济学的演化脉络来看,他们将贫困问题作为发展经济学研究的核心使得发展经济学研究的视野过于狭窄。在以亚当·斯密等人为代表的古典经济学家看来,经济发展是劳动分工与市场演进相互作用的过程,涉及专业化、市场网络、工业化、企业、贸易、技术进步、城市化等一系列广泛的相关议题(杨小凯,2003),贫困问题解决是经济发展的题中应有之义。现在发展经济学不关心分工和市场演进相互作用的过程,转而考虑贸易保护政策、产业政策等一系列国家主导的经济发展战略,这实际上考虑的是给定资源最优配置的问题,是对经济发展问题研究视野的狭窄化。随着发展经济学在政策实践中的失败,发展经济学家转而直接关注贫困问题,则使得经济发展问题研究的

视野进一步狭窄化。这反过来对贫困问题的根本解决可能产生有害的影响。反观我国扶贫工作的巨大成就,虽然与大量扶贫工作的进展有密切的关系,但是也离不开劳动分工与市场演进相互作用的过程。贫困问题的根本解决需要将广大的贫困人口纳入到社会分工中来,而这个过程不可避免地会带来新的问题。这些问题正是发展经济学家应该关注的议题。总之,发展经济学应该回归和重视经济发展机制的研究。

本文剩余部分安排如下:第二部分简要回顾他们的学术生涯及其主要论著;第三部分对因果关系与实地实验研究方法及其在经济学研究、经济政策制定和企业经营管理实践中的价值进行述评;第四部分重点对他们贫困问题所展开的实地实验及其结论进行扼要总结,并且站在发展经济学演化脉络和我国的扶贫实践对其进行评价;最后是结语。

二、学术生涯与主要论著概述^①

阿比吉特·班纳吉(Abhijit Banerjee),1961年出生于印度孟买。他先后在加尔各答大学和贾瓦哈拉尔·尼赫鲁大学分别获得理学学士和硕士学位。1988年在哈佛大学获得博士学位。其后,班纳吉曾任教于哈佛大学、普林斯顿大学和麻省理工学院。2003年至今,担任麻省理工学院福特基金会国际经济学教授和阿卜杜勒·拉蒂夫·贾米尔贫困行动实验室(Abdul Latif Jameel, J-PAL)主任。班纳吉曾担任过经济分析与发展研究局(BREAD)主席、国家经济研究局(NBER)研究员、基尔研究所国际研究员等,还任联合国2015年后发展议程高级别小组、世界银行研究评估小组、全球发展中心高级市场承诺工作组等国际性组织的成员。班纳吉还曾担任《美国经济评论》编委会成员、《经济学季刊》副主编,并且是众多经济学期刊的审稿人。

班纳吉的研究主要集中在经济发展、信息经济学、收入分配理论和宏观经济学领域。1988年,班纳吉在马斯金(Eric Maskin)的指导下从事理论研究,以《信息经济学论文集》作为学位论文取得博士学位。此后,他采用信息经济学的分析工具探讨了各种纷繁复杂的经济问题,对这些问题的探索正好提升了他应用信息经济学分析工具的能力,同时也让他找到对收入分配理论这一新的研究兴趣。对该领域的研究至少可以追溯到他对风险分摊与收入分配理论的研究(1991);他对不完全信息与永久收入假说的探讨(1990)则使其研究延伸到了宏观经济学领域;更为重要的是,他将信息经济学应用于发展中国家经济发展问题的研究中。1993年的《职业选择与经济发展过程》和1994年的《贫困、激励和发展》两篇论文标志着班纳吉向发展经济学研究转型之路的开始。随着他对发展中国家特别是贫困问题的研究的深入,为了验证具体贫困解决政策的有效性,他的研究方向从理论研究转向实证研究。由此,班纳吉走上了发展经济学实证研究之路,并发表了多部著作和大量论文。

他对发展经济学和贫困问题以及其他经济问题的贡献给他带来了许多荣誉和奖项,包括印度马哈拉诺比斯纪念章(2000)、美国政治科学协会迈克尔·沃勒斯坦奖(2006)、BBVA基金会“知识前沿”发展合作奖(2009)、Infosys社会科学奖(2009)、加巴隆国际经济奖(2013)、阿尔伯特·赫希曼奖(2014)等。此外,他还入选《外交政策》杂志2011年全球100位顶尖思想家。

埃丝特·迪弗洛(Esther Duflo),1972年生于法国巴黎,是一位杰出的女性经济学家。她在法国高等师范学院主修历史和经济学,在获得经济学硕士后来到麻省理工学院,并于1999年获得博士学位。毕业之后,迪弗洛一直任教于麻省理工学院,2005年至今担任麻省理工学院Abdul Latif Jameel扶贫与发展教授和Abdul Latif Jameel扶贫行动实验室(J-PAL)联合主任。她还是一些重要机构或组织的领导或成员,包括经济政策研究中心(CEPR)发展项目主任、国家经济研究局(NBER)副研究员、经济分析与发展研究局(BREAD)董事、瑞银中心顾问委员会成员、全

^①由于相关论文实在太多,这里仅列出有代表性的著作。

球投资基金(GIF)董事会成员(2014—2017)、总统全球发展理事会成员(2012—2017)。此外,她现任或曾任许多顶级经济学期刊包括《美国经济评论》、《发展经济学期刊》、《经济展望期刊》等的主编或编辑。

迪弗洛早期的兴趣是历史,在巴黎高等师范学校得到托马斯·皮克提(Thomas Piketty)的指点而转向经济学。在麻省理工学院攻读博士期间,她师从班纳吉,在班纳吉的指导下以3篇发展经济学的实证论文作为学位论文获得博士学位。她的博士论文荣获法国科学院博士论文奖学金以及阿尔弗雷德·斯隆博士论文奖学金。迪弗洛的研究一直围绕发展经济学特别是贫困问题展开。她力图了解穷人生活的方方面面,涉及预防保健、教育、小微金融、环境以及性别和政治等问题,在这些方面出版和发表了许多著作和论文,包括《与贫困作斗争》(卷I和卷II,2010)等。

迪弗洛因其对贫困问题的深入研究被授予了一系列荣誉和奖项,包括法国最佳青年经济学家奖(2005)、BBVA基金会发展合作知识前沿奖(2009)、Infosys社会科学奖(2009)、麦克阿瑟“天才奖”获(2009)、约翰·贝茨·克拉克奖章(2010)、哈佛肯尼迪学院托马斯·谢林奖(2011)等。

值得一提的是,班纳吉和迪弗洛是师生和夫妻双重搭档。他们携手对发展经济学和贫困问题进行了深入的研究,出版了《贫困的本质:我们为什么摆脱不了贫穷》(阿比吉特·班纳吉和埃斯特·迪弗洛,2018),从各个层面探究导致贫穷的根源。该书获得2011年度英国《金融时报》和高盛商业年度最佳图书奖,是了解贫困问题的一本必读书。他们为了评估实地实验的进展,总结了一套实地实验研究方法,编辑出版了两卷本:《实地实验手册(卷I和卷II)》(Banerjee和Duflo,2017)。该书在提供了实验研究策略的同时,还提供了一份关于最新研究成果和有待探索领域的综合目录,是实地实验领域研究人员必备的参考书。2003年,班纳吉和迪弗洛携手塞德希尔·穆来纳森(Sendhil Mullainathan)一起在MIT创建了Abdul Latif Jameel贫困行动实验室(J-PAL)旨在确保基于科学依据来制定扶贫政策。在他们的指导下,该实验室已发展成为一个致力于通过随机控制实地实验来解答重要政策问题,以及对抗贫困问题的研究员的全球性网络,其研究成果已赢得国际社会的认可。

迈克尔·克雷默(Michael Kremer)出生于1964年。1985年,克雷默以全美优等生联谊会成员的身份在哈佛学院获得社会学学士学位,之后在肯尼亚卡卡梅加区埃希斯鲁中学担任教师兼行政管理人员(1985-1986),期间组建并担任非盈利组织——世界教育(WorldTeach)的执行董事(1986—1989)和主席(1990—2017)在发展中国家安置志愿教师。1992年在哈佛大学获得博士学位之后,辗转任教于麻省理工学院、芝加哥大学和哈佛大学。1999年起任哈佛大学经济学教授。2003年至今任哈佛大学经济系发展中社会盖茨讲席教授。他还曾担任国家经济研究局(NBER)副研究员(1999年至今)、美国国际开发署发展创新创业科学总监(2010年至今)、世界银行发展经济学研究组顾问(2007—2016)等,并曾担任《发展经济学期刊》和《经济学季刊》等期刊的副主编。此外,他还是经济分析与发展研究局(BREAD)联合创始人,在2001—2004年期间组织了多场发展经济学会议。

1992年,克雷默在哈佛大学师从马斯金从事理论研究,以优异的表现取得博士学位,他的博士学位论文荣获哈佛大学威尔斯经济学论文奖。随后,他发表了与经济发展和增长相关的论文,提出了经济发展的“O”形环理论(1993)。他还探讨了人口增长与技术变迁问题(1993)以及储蓄与经济增长(1994)等问题。如果说他这一阶段的研究主要是围绕理论问题展开,那么随后的研究则紧密地将理论和实践相结合。肯尼亚的经历使他敏锐地发现了发展中国家的教育问题。在对教育问题的研究过程中,他又发现了贫困人群驱虫问题的重要性,由此更扩展到发展中国家的卫生保健的相关问题,主要涉及驱虫药的使用、传染病与新药、艾滋病的预防与控制、饮用水等。在这些研究的基础上,他还帮助制定了新药相关政策,以此刺激发展中国家在新药

研究及销售方面的私人投资。这些研究也使他关注了处在贫困中的人群的心理问题。他和迪弗洛等人合作研究了农业中的化肥使用状况,借助行为经济学理论设计了有效的助推农民使用化肥的策略,并且在总结贫困心理问题的基础上,提出了行为发展经济学理论。

克雷默对发展中国家贫困问题的研究也给他带来了一系列的荣誉和奖项,包括科学家和工程师总统早期职业奖(1996)、麦克阿瑟“天才奖”(1997)、国际卫生经济学协会授予的Kenneth J. Arrow健康经济学最佳论文奖(2004)、拉丁美洲和加勒比经济协会的社会政策最佳论文的胡安·路易斯·隆多尼奥奖(2014)等。2004年,他与雷切尔·格伦内斯特合著出版了《妙手良方:为被忽视疾病的药物研究创造激励》一书,被美国出版商协会授予强效医学领域最佳专业学术书籍奖。

简要回顾三位诺贝尔经济学奖得主的学术研究之路,可以看出他们或者早期从事经济学的理论研究而后转向实证研究,或者从一开始就主要从事经济学的实证研究。这是当今经济学研究发展的一个重要趋势。20世纪90年代末期以来,在迪顿(Deaton)等人的推动下,高质量的微观经济数据库得以建立,并且计算机计算能力有了大幅度提高,发展经济学也经历了从理论研究向实证研究的重大转变。班纳吉、迪弗洛和克雷默不仅顺应了这个趋势,而且将实地实验引入到发展经济学的实证研究之中,大大提高了贫困治理机制及其结果之间因果关系的可信度,弥补了工具变量法、双重差分法和回归断点设计等因果关系可信性革命中传统分析方法的不足之处。

三、因果关系与实地实验

如前所述,解决贫困问题的关键在于了解处于贫困中的人们行事背后的动机,其实质就是要通过实证研究搞清楚贫困背后的因果关系。然而,在经济学实证研究中已经有许多因果关系实证研究可信的分析工具,为什么实地实验方法会成为发展经济学领域的主要研究方法?实地实验是如何具体展开?它具有哪些优点以及存在哪些问题?我们在简要回顾统计推断基本原理的基础上,对比经济学与自然科学因果关系研究的异同,并且逐一探析上述问题。

(一)因果关系与统计推断的基本原理

自然科学和社会科学对因果关系的分析是通过数据对统计推断来进行的。统计推断的基本原理可以归纳为判断数据与基准之间的偏差是否大到随机性无法解释的程度。这可以用检验硬币是否均匀的例子对此进行解释^①。为了检验硬币是否均匀,假设研究者将一枚硬币掷100次,如果硬币是均匀的,我们知道它将差不多50次正面朝上,50次反面朝上。这就是随机现象中的规律性。然而,这种规律性也表现出某种随机性。如果你再掷100次硬币,正面朝上的次数不会和之前的一样,可能更多也可能更少,通常正面朝上的次数不会准确等于50次,而是与50次有一定的偏差。通常情况下,这种偏差可以由数据本身的随机性来解释。然而,有时偏差会大到一定的程度,以至于无法由数据本身的随机性来解释。例如,考虑两个极端情况,100次都是正面朝上,或者0次正面朝上。那么,这种情况离硬币是均匀这个理论下可能出现的结果偏差太大,无法用随机性来解释。此时,我们就推断硬币是均匀的这个理论可能是错误的。

当然,要推断理论是否错误,不需要这么极端的情况,只要偏差大到随机性无法解释的程度,就能有把握推断理论是错误的。这个因果系统统计推断的基本原理不管对自然科学还是社会科学都是一样的,但是由于学科性质的不同和条件的限制,自然科学和社会科学对基准的选

^①这个例子来自一个历史典故。据说北宋时期,将军狄青为了提高士气,想出了一个十分巧妙的方法。他设下祭坛参拜神明,在祈祷的时候,随手拿出100个铜钱,口中立下誓言:“这次和叛军作战生死未卜,若能取得胜利,就请神灵保佑,让这100枚铜钱全部正面朝上!”神奇的一幕发生了,这100枚铜钱落在地上之后,士兵们前来围观,发现这100枚铜钱全部正面朝上,便以为是神灵保佑,军队之中顿时爆发出阵阵欢呼,宋军顿时士气大振。

择有大的不同,因果关系推断的有效性争议由此而起。

(二)自然科学中的随机控制实验

自然科学中因果关系统计推断中的基准可以通过随机控制实验来设定。以新药有效性检验为例,在新药推出之前,为了要正确地检验新药的效果,研究人员通常使用随机控制的双盲实验来排除其他因素的影响。随机双盲实验的关键有两个:第一,将实验对象随机分成实验组和参照组,随机的目的在于排除两组人之间的系统性差异,避免出现诸如一组志愿者男性的比例特别高,而另一组女性的比例特别高的情况。否则,将无法搞清楚感染人数减少到底是因为新药起作用的缘故,还是男女性别体质差异的缘故;第二,要给对照组注射安慰剂,安慰剂外观要跟新药一模一样,除了研究人员之外,注射人员与实验组和参照组的志愿者彼此都不知道。之所以这么做,是要排除人体自身免疫力以及心理作用的影响。只有经过如此细致的实验,才能将参照组作为药品即使无效仍然会好起来的人数作为基准,而将实验组与参照组之间的差异归因为药品的效果。

随机控制实验背后有两种看待因果关系的不同视角。一种视角是“保持其他条件不变”的视角,即研究人员通过随机双盲实验,确保实验组和对照组除了注射新药或是安慰剂之外,其他各个方面是一样的。另一种视角是“反事实”的视角。即要检验新药的效果,就要找到假设新药无效的情况下人们感染传染病的人数。参照组要回答的正是这个假想的问题,计量经济学家称之为“反事实”。而经济学或者说社会科学因果关系的检验通常难以通过实验来设定基准,而是依赖于被动收集的非实验数据。这些数据是各种因素共同作用的结果,为了厘清众多复杂的因素中的因果关系,计量经济学家沿着这上述两个视角设计和发展了各种实证分析工具。以教育回报率的实证研究为例,这个问题要分析的教育程度这个“因”对工资这个“果”的影响。

(三)因果关系问题的计量经济分析工具

要正确地分析教育对工资的影响,我们需要排除其他因素的影响,或者说保持其他因素不变的情况下,探讨教育程度的变化对工资的影响,这实际上是从“保持其他条件不变”的视角来分析因果关系问题。为了保持其他因素不变,早期的计量经济学分析借助的工具是多元线性回归模型。计量经济学家通常将除了教育之外的其他影响工资的因素分为两类:一类是诸如工作经历容易度量的因素,称之为控制变量;另一类是诸如能力难以度量或者无法观测的因素,称之为误差项。在经典的多元线性回归模型中,将工资这个被解释变量看成解释变量即教育程度,控制变量即工资、经历等可以观测的因素以及无法观察到的随机误差项的函数。在满足线性回归模型的假设条件下,探讨在保持控制变量不变的情况下解释变量对被解释变量的影响。这些线性回归模型假设正是自然科学中随机控制实验要满足的基本要求。因此,可以把满足线性回归模型假设条件的多元线性回归模型看成社会科学对自然科学随机控制实验的模拟。

由于控制变量已得到明确控制,多元线性回归模型对因果关系分析的关键就在于误差项问题。在线性回归模型条件中,最重要的假设是零条件均值假设^①,即给定解释变量和控制变量的条件下,随机误差项的均值为零。零条件均值假设也被称为严格外生假设。在教育的回报率的例子中,也就是说教育程度和工作经历等其他控制变量外生于随机误差项,就像在新药检验的例子中注射新药或安慰剂,这是由研究人员外生决定的,而不是由实验对象自己选择的。

然而,在社会科学研究中,因果关系分析中所采用的非实验观测数据往往是人们选择的结果,即数据具有内生性。因此,严格外生的假设并不成立。以教育回报率为例,通常将无法观测的能力因素纳入随机误差项。但是,在这一模型中,教育程度并不是外生的,而是与能力因素有

^①经典线性回归模型包括六个假设,即参数的线性、随机抽样、无完全多重共线性、零条件均值、同方差、正态性等假设,其中前面五个假设也称为高斯—马尔科夫假设,在这五个假设下,使用普通最小二乘法可以得到参数的最优线性无偏估计量。

关。一个很大的可能是,能力越高的人会选择更高的教育程度,而能力越高的人原本也就更有可能获得高工资。因而,我们无法区分出一个人工资高,到底是因为受教育程度高导致的,还是其能力高导致的结果。假设教育程度是外生的,那么使用多元线性回归模型就很可能高估教育的回报率。

因此,基于多元线性回归的分析更多地被看成是一种相关关系分析,而不是因果关系分析。计量经济学家转而借助“反事实的”视角对因果关系进行探讨,设计出了其他能够估计因果关系的可信方法,称之为计量经济学的可信性革命(credibility revolution)(Angrist和Pischke, 2010),已被广泛地应用于各种微观实证分析之中的微观计量方法包括工具变量法、双重差分、回归断点设计等。计量经济学的可信性革命的要点是找到所谓的“偶然实验(accidental experiments)”,即现实世界中刚好发生的某种与随机控制实验性质类似的事件。以采用断点回归设计估计高等教育的回报率为例。假如想要估计高等教育的回报率,若采用多元线性回归模型,可能会高估大学教育的回报率。为此,可借助高考切分线(设分数线为550分)作为断点,选择551分考上大学的一组人作为实验组,选择449分没有考上大学的另一组人为对照组。由于分数线的划定具有一定的随机性,这两组人之间1分之差代表不了某种系统性的差异。因此,它在很大程度上类似于新药有效性检验里面的实验组和对照组,这使得研究人员可以有比较大的信心地将两组人之间的差异归因为大学教育的影响。

(四)实地实验方法

计量经济学的“可信性革命”提高了许多经济问题因果关系分析的可信度。但是,其研究却大大地受制于偶然实地的可得性。如果研究人员借助已发生事件对已有的数据进行分析,这必然大大地限制研究人员想要探讨的问题的范围。发展经济学家要探讨的贫困问题解决机制及其后果之间的因果关系。许多机制是他们根据贫困问题的具体情况的分析,结合经济学理论设计出来的,对于这些机制是否能切实有效地发挥作用,也就是对其因果关系的检验,实际上没有相关的数据。即使有,这些数据也因为内生性问题,无法对因果关系进行有效的检验。这正是班纳吉、迪弗洛和克雷默要在发展经济学实证研究中大量采用实地实验研究方法的主要原因。

为了对贫困问题解决机制是否有效的因果关系问题进行可信的推断,发展经济学家采用将自然科学的因果关系研究的实验研究规范引入到其实证研究中,在现实世界中进行经济学的随机控制实验,这正是我们称之为实地实验的一个原因。发展经济学家按照实验的规范,在现实世界中对所要检验的问题进行设定,将实验对象进行随机分成实验组和参照组分组,然后开展随机控制实验。由于除了干预不同之外,发展经济学家通过随机分组尽量控制其他影响因素使其保持不变,因而实验组与参照组的不同表现可以归因为干预的不同。然而,与自然科学的实验不同,实地实验通常要探讨人们对方案的反应。因此,或许实验对象不知道自己正在参与实验,但是他们通常可以看到相关的政策干预。也就是说,它要在现实世界的环境中理解对所设计的方案的反应,这不同于新药检验中劳动双盲实验,这是实地实验之所以称成为实地的另一个原因。

从本质上看,实地实验是回复到自然科学的随机控制实验的做法。这种方法有许多优点。首先,研究人员可以使用实地实验将实验对象进行随机分组,从而有效地消除选择偏差问题,从而有效地避免传统计量经济方法特别是多元线性回归方法对外生性假设的争议。其次,实地实验使得研究者化被动为主动,而不再局限于对已经发生过的时间进行评估,研究人员可以根据已有的经济学理论,设计出他们认为可能有效的机制,然后通过实地实验的方法对这些机制进行检验,实地实验领域的最重要的局限就是实验设计者的想象力不够丰富。再次,实地实验提供了不断改进的空间,传统的计量分析方法通常只能对某个问题进行一次性的分析,而实地实验则允许在前实验的基础上进行后续的实验,这使得研究人员可以展开迭代研究过程,既可

以从先前的实验中发现问题的,又可能从中观察人们对政策变化的学习效应。最后,实地实验的规划和实施迫使研究人员要实地接触并花时间了解当地的实际情况,经常与政府或非政府机构密切合作,这一方面有助于研究人员了解机制设计的各种约束条件,从而发现经济学书本上未加以考虑的问题进而发现新的机制,另一方面,反过来也有助于经济学理论的完善,从而提高我们设计更好政策的能力。这些正是他们对实地实验方法的一个主要贡献。

由此,实地实验有着其他工具所不具备的优势,可以使得经济学家主动地对经济问题背后的因果关系进行主动的设定,从而有效地解决困扰传统多元线性回归模型经济学家的内生性问题,同时使得经济学家不再受制于偶然实验的可得性,从而开辟经济学因果关系检验的新空间。在这个意义上,我们认为,就像一个优秀的厨师一样,他的工具箱里面有不同的刀具,有的适合砍骨头,有的适合切肉,经济学家应该重新审视实地实验的方法,将其纳入因果关系检验的工具箱中。因果关系工具箱中的不同工具有着不同的功效。

实地实验作为因果关系检验的有效工具,也是政策制定的有效工具。班纳吉、迪弗洛和克雷默的实地实验,实际上也是对解决贫困问题的相关政策制定进行试点,通过在小范围内对政策进行实地实验,获得相关的数据,对政策的有效性进行分析,找出有效的政策然后加以推广。我国的经济改革一直以来都秉持“摸着石头过河”的哲学,从家庭联产承包责任制、乡镇企业发展、国有企业改革、经济特区建设到区域金融改革、教育体制改革以及环境保护等政策的实施,通常都是通过政策试点来寻找有效的政策然后加以推广。在这个意义上,我国的经济改革可以说是一场史无前例的大规模实地实验。企业经营管理也是如此,以往企业管理者往往依靠直觉或者从实践中总结出来的经验对企业进行管理,企业的管理者如果学会并善于使用实地实验的方式,就可以利用实地实验的方式来对企业的经营管理策略进行测试,通过实地实验的方式对产品定价、产品和服务创新进行检验,从而发现最有效的经营管理方式。

当然,实地实验并非完美无缺的。人们对实地实验的最大质疑是,从实地实验所得到的因果关系结论是否具有一般性,特别是大规模加以推广是否还能达到预期的效果。这实际上是对实地实验外部有效性的质疑,这也是自然科学和社会科学面临的共同问题。回到新药有效性检验的例子,新药的开发生通常也不是一开始就在人身上做实验,而是选择小白鼠做实验,因此,也面临着从小白鼠身上得到结论是否能直接推广到人身上的问题。此外,由于病毒的抗药性问题,新药也可能随着时间的推移而失效,这些都是外部有效性的问题。当然,由于自然科学的性质,外部有效性的质疑通过对实验的再一次重复通常就可以得到解决。然而,社会科学因果关系的外部有效性问题则有着特别的难处,传统的计量经济学工具在对因果关系进行实证分析之后,对结论的有效性进行稳健性检验,来回答外部有效性的质疑。而实地实验降低对具体问题因果关系分析的难度,却增加了外部有效性的质疑。解决实地实验方法的外部有效性问题是他们对实地实验方法的另一个主要贡献。接下来,我们将对贫困问题的实地实验进行总结,阐述他们如何解决因果关系推断以及外部有效性的问题。

四、新实证发展经济学的贫困治理之道

在迪顿等人的推动下,发展经济学家逐渐将发展中国家的贫困问题从宏观的视野转向关注微观的问题。在他们看来,发展中国家的贫困问题是多方面的,涉及教育、卫生保健、信贷、政治以及心理等。他们对这些小问题的研究取得了丰硕的成果,同时通过实地实验的改进然后再实验等辛苦的工作,对其结论进行推广,积极回应了外部有效性的质疑。这里首先阐述他们对贫困问题的实地实验的主要研究成果,然后在发展经济学的演化脉络中评价其得失。

(一)发展经济学微观化与实地实验

从20世纪90年代中期开始,克雷默及合作者与非政府组织合作,在肯尼亚西部就如何提高

教育质量展开了一系列实地实验。随后,班纳吉和迪弗洛加入进来,他们或相互合作或与其他研究人员合作,将实地实验扩展到了卫生保健、信贷、政治以及心理等一系列相关的问题。研究的结果表明,穷人之所以这样或那样行事,与私人部门和公共部门的效率低下问题密切的关系。为此,他们针对效率低下问题设计了许多激励机制,并且通过实地实验及其他方法,为最有效地解决效率低下问题提供科学证据^①。

1.提高教育质量

发展中国家发生贫困的一个根源是教育问题。尽管众多的发展经济学家强调人力资本在经济发展中的重要作用,也确信教育是形成人力资本的最重要方式。许多发展中国家的教育依然泛善可陈。克雷默早期的研究是教育投入不足问题,例如发展中国家诸如缺乏必要的学习材料、儿童经常生病旷课或者吃不饱饭导致健康问题普遍存在、教师不够用心等等。克雷默及其合作者构想了许多干预方案,包括提供教科书和挂图等学习材料、对儿童进行驱虫以及提供膳食计划的方式对儿童教育进行干预,并且将教师的奖励与学生的学习成绩挂钩以激励教师。与经济理论的预期不一致的是,实地实验的结果表明,提供更多的资源对教育质量的影响是有限的,例如,提供更多的教科书并没有提高平均考试成绩,但确有提高最有能力学生的考试成绩。也有与经济理论预期一致的,例如将教师的激励与学生成绩挂钩提高了与激励措施相关的考试成绩但与激励不相关的考试成绩则不受影响。这个结果与多任务委托—代理理论的预期一致(Kremer,2003)。

增加教科书等教育资源的投入没有起到应有的作用,可能是因为额外的教育资源投入只有在满足特定条件时才能发挥作用。在许多发展中国家,教学方式和课程设置都不能适应大量儿童涌入初等教育的情况。因此,他们将研究的焦点转向教学改革,即考虑如何使教学更好地与学生的学习水平相匹配。他们在印度展开了一项针对弱势学生学习的两种干预措施的实地实验(Duflo等,2011)。一个干预措施是补习项目。他们雇用人员在课外为表现不佳的三年级和四年级学生提供补习,另一个是计算机辅助学习项目,让四年级的孩子每周在一台共享计算机上玩两个小时的数学拼图游戏。实地实验发现,与单纯增加教育资源投入的研究相反,在一年和两年之后,上述的这两种干预措施对学生的成绩都产生了显著的积极作用。

他们还发现,在发展中国家教师缺勤非常严重。他们与一个非政府组织合作,从其开办的只有一个老师的学校中随机选取了一些学校,通过摄像机记录这些学校的教师出勤情况,并向出勤情况好的教师发放额外的奖金。实验发现,与对照组学校相比,实验组学校的教师缺勤率下降了一半,学生的学习也有所改善(Duflo等,2012)。更重要的是,他们对教师聘任机制进行实地实验。借助肯尼亚增加教师聘任数量以降低班级规模的契机,他们在正常的教育部公务员制的教师渠道之外,增加聘任一些以年度合同聘用的、工资通常低于公务员制教师的短期合同制教师,以此来了解短期合同制教师与公务员制教师对学生学习的影响。实地实验发现,一方面,接受短期合同制的教师缺勤率较低,学生的考试成绩也有显著提高;另一方面,在接受短期合同制教师资助的学校中,公务员制教师的缺勤率反而上升了,同时,即使班级规模消减了,被随机分配留在原来班级仍由公务员制教师教的学生,成绩并没有显著提高(Duflo等,2015)。

2.预防保健问题

贫困人口的健康投资不足是困扰发展中国家人力资本投资的另外一个问题。为什么会这样?如何加以改善?他们根据理性的人力资本投资模型,预期人们只有在私人边际收益大于边际成本时,才会进行投资。然而,许多健康方面的投资具有很强的正外部性,这使得预期私人边际收益小于社会边际收益,由此导致消费者对预防保健的投资不足。这种正的外部性对发展中

^①他们进行了许许多多的实地实验,这里择其要者进行论述。

国家传染病的预防具有重要意义。然而,外部性大小的实证研究却相对缺乏。为此,克雷默及其合作者巧妙地设计了一个实地实验用于估计驱虫药的直接影响和外部性。他们先按地理位置对75所实验组小学进行分组,然后按地点的字母顺序进行分组,这一设计使得驱虫药的效果取决于特定地理距离内当地学校人口的总密度,而且使得分配给实验组小学的儿童人数应与其他当地可观察和不可观察的因素不相关。他们借此估计外部性的大小,发现外部性对蠕虫感染率以及随后的学生上课的出席率有显著的影响,这种影响可以延伸到距离实验组小学至少3公里的地方(Miguel和Kremer,2004)。

定价是另一个影响因素。定价决定了使用者要付出的代价,对药品使用的影响显而易见。预防保健服务是否应该收费、收多少费,这是长期以来相关政策考虑的热点问题。反对收费的人认为,收费会阻碍穷人使用卫生保健服务,即使从成本效益分析的角度看是合算的;支持收费的人则认为,收费会让人们更加重视产品的使用。而问题的关键在于,穷人对定价的敏感性。克雷默及其合作者最先对价格如何影响卫生保健产品使用展开实地实验评估。在上述驱虫药外部性的实地使用中,他们从参加上述驱虫药实地实验的75所小学中随机挑选了25所,让这些学校的学生家长必须为孩子分摊驱虫药的费用(Miguel和Kremer,2004)。结果发现,收费大大降低了驱虫药的使用,表明穷人对卫生保健产品的需求对价格非常敏感。这一研究推动了发展中国家为卫生保健提供了大量的补贴。

在发展中国家中,由于存在医疗卫生服务质量差、护理人员缺勤率高等问题,无形中也提高了穷人使用相关服务的成本,从而降低了使用效率。他们发现,印度农村地区只有2%的1—2岁儿童接受了所推荐的基本疫苗接种。如此低的接种率的一个重要原因是,负责疫苗接种的卫生工作人员经常不在岗。为了提高印度农村地区的疫苗接种率,他们设计了流动疫苗接种诊所(Banerjee等,2010)。并保证卫生工作人员总是在场,同时还随机选择一些流动疫苗接种诊所,对接受疫苗接种的家庭提供小额实物奖励。实地实验结果发现,参照组的疫苗接种率为6%,流动诊所中没有提供奖励的为18%,流动诊所中提供奖励的为39%。表明这些激励措施对鼓励家庭坚持疫苗接种特别有效。

3. 小额信贷

除了人力资本投资之外,穷人在实物资本上的投资同样乏善可陈。长期以来,发展经济学家注意到贫困国家与富裕国家之间的平均生产力存在巨大差异的事实,却忽视了穷国内部生产力也存在巨大差异的问题。正如班纳吉和迪弗洛所指出的,贫困国家和富裕国家之间的生产力差异很大程度上取决于低收入国家内部的生产力差异,在贫困国家内部,有些人或公司使用了最新的技术,而另一些从事类似生产活动的人——通常是穷人——却使用着过时的技术(Banerjee和Duflo,2005)。那么,为什么这些穷人不使用先进的技术?一个可能的原因是存在信贷约束,贫困者得不到信贷的支持,被迫使用过时的生产技术。

为了缓解穷人的信贷约束,许多国家和非政府组织都极力推动开展小额信贷运动。但是,有关小额信贷运动效果的研究相对缺乏。班纳吉和迪弗洛进行相关的实地实验研究。他们对印度海得拉巴市标准团体小额信贷进行实地实验评估。这个评估针对的是有可能成为企业家的女性,跟踪时间长达3.5年。他们对小额信贷的影响进行短期和中期两次评估。短期评估的时点是12—18个月后,得到如下结论:第一,小额信贷确实使部分家庭借到了更多钱,但是总体比例并不高,只有约1/4的满足条件的家庭能从小额信贷机构获得借款;第二,一些通过小额信贷机构借款的家庭使用贷款替代非正规贷款,小额信贷贷款增加而非正规贷款减少,因而借款总额并没有显著上升。这两个证据表明,小额信贷的需求并不是很高。第三,通过对人均消费支出等指标的分析也未发现小额信贷带来了显著差异;第四,有机会获得小额信贷的妇女成为企业家

的可能性也没有显著增加,小额信贷对妇女拥有的企业的平均盈利能力有显著影响,但是似乎更有助于最赚钱的企业。中期评估的时点是3.5年,这时对照组也获得了小额信贷,但实验组家庭能够获得借款的时间要长得多,然而并未发现两组之间存在显著的差异。综合来看,没有证据能支持小额信贷作为经济发展的主要动力的观点(Banerjee等,2015)。

4.心理偏差

小额信贷未起到应有的作用,这难免使人们对穷人的行为是否理性产生怀疑。早期的发展经济学家假设穷人是理性的,进而探讨了贫困对处于贫困中的决策者产生的约束。随着行为经济学的发展,发展经济学家引入行为经济学的相关理论反思理性选择模型,从心理学的角度对贫困问题进行重新思考。根据行为经济学的理论,理性人假设与现实中人们的行为并不相符。现实中的人们在决策过程中存在许多非理性的特征,塞勒(Thaler)将这些特征归纳为有限认知、有限自利和有限自制力。其中,对实物投资不足最有解释力的是有限自制力问题,它导致人们缺乏自我控制因而无法坚持所制定的计划,这也称为现时偏差或者双曲线贴现率。

他们借助行为经济学理论展开了一系列实验,用于探讨困扰许多小农特别是非洲撒哈拉以南地区小农的一个大难题:为什么诸如化肥等简单的现代农业生产技术,虽然能够带来很高的回报率却没有得到广泛的应用?这实际上是一个实物资本投资的问题。他们采用有限自制力的假设构建了一个模型来解释农民的行为。在他们的模型中,有些农民存在现时偏差(或者双曲线贴现率),这意味着打算购买化肥的农民即使有钱也会推迟购买,直到最后期限到来。但是,到那时他可能因为没有钱而无法购买,导致化肥的投入不足。在这个模型的基础上,他们提出并比较了两种替代性的政策干预措施:相对较大的补贴和在农民有钱时为其购买肥料提供限时使用的小额折扣。按照行为经济学理论,限时小额折扣比大额补贴更能抵消现时偏好的影响。借助实地实验,他们评估了这两种干预措施的效果。结果表明,农民在提供限时小额折扣(以免费送货的形式)时多购买了50%的肥料。其效果比在本季晚些时候提供免费送货和较大额的化肥补贴的效果更大,并且与理论预期一致(Duflo等,2011)。

这项研究是将行为经济学融合到发展经济学研究方面的一个典范,它促进了发展经济学对穷人认知和心理决策的研究。克雷默在总结将行为经济学运用到发展经济学的研究经验的基础上,提出了行为发展经济学,从行为经济学的视角对发展经济学的诸多问题重新展开探讨(Kremer等,2019)。

5.性别与政治

在理解经济发展和贫困问题时,必须考虑发展中国家的性别与政治制度这一重要因素。迪弗洛与合作者借助印度政治改革的契机,评估了政治领导人性别对政策选择的影响。1993年,印度联邦政府出台了一项新的宪法,规定每个邦必须保留三分之一村委会主席职位给女性领导人。由于村委会在当地基础设施投资中发挥了很大的作用,这为研究提高印度妇女政治地位的政治改革提供了一个巧妙的偶然实验。迪弗洛及其合作者借此估计随机挑选的女性领导人的影响。实证结果表明,女性领导人的决策似乎更符合女性的偏好。在西孟加拉邦,农村妇女更关心饮用水和道路,而农村男子更关心教育;西孟加拉邦的女性领导人确实比男性领导人在饮用水和道路上投入更多,而牺牲了教育;在拉贾斯坦邦也有类似的情况,农村妇女比农村男子更关注水资源,但对道路的关注度较低,女性领导人在水资源投资方面的支出也高于道路(Chattopadhyay和Duflo,2004)。

在后续的研究中,迪弗洛及其合作者利用了同样的基础自然实验,但是增加了来自约500个村庄的样本的实地实验数据作为补充。他们发现,一个村委会领导职位的一再“保留”给女性领导人大大提高了未来选举中女性候选人获选的可能性。他们还发现了这一结果背后的

一个重要机制,即村委会领导职位一再“保留”给女性领导人减少了选民对女性领导人的成见,特别是减少对妇女不能成为有效决策者的偏见(Duflo等人,2009)。

他们对女性政治领袖的研究拓展了发展经济学的范围,促进了发展中国家和发达国家的性别和政治的相关研究。当然,给女性赋权会影响经济发展,反过来,经济学发展也会影响对女性的赋权,实际上可能存在双向因果关系,这是有待探讨的问题。

(二)外部有效性与实地实验结论的推广

班纳吉、迪弗洛和克雷默等人及其合作者展开的实地实验取得了丰硕的成果,同时也受到许多质疑,人们怀疑他们的实地实验所得到的结论在进行推广时是否还成立。发展经济学领域的许多实地实验通常由一个管理良好的非政府组织在相对较小的范围内展开,如果将实地实验所得的结论以政策的方式进行大规模的推广,有许多因素可能会导致结论不再可靠。归纳起来,这些影响实地实验外部有效性的因素有五类:均衡效应、溢出效应、背景依赖、随机化偏差和试点偏差。均衡效应和溢出效应与规模扩大直接相关,小规模实地实验可以由局部均衡模型进行解释,但是大规模的推广可能会影响到工资和价格等变量;背景依赖指的是实地实验可能与文化、制度等背景相关,推广的时候可能因为这些背景的不同而发生改变;随机化偏差指的是,同意参加小型实验的受试者可能与其他人群不同;试验偏差则是指,需要高度监控的较小项目的结果可能无法在大规模运行的项目中复制,以实地实验为基础的政策若要大规模推广,势必会对政府的组织管理能力提出挑战。

围绕实地实验的外部有效性,班纳吉、迪弗洛和克雷默与实地实验方法的质疑者展开了激烈的争论。他们在许多论文中讨论和分析了对外部有效性的挑战,探讨了应对、处理或者减轻外部性挑战的方法。有些外部性有效性问题可以通过实验的复制来解决,然而还有更多的实地实验的结论难以通过复制实验来解决。为此,他们对已有的实地实验进行改进,重新实施新一轮的实地实验,以此回应质疑,并且推广其实地实验的结论(Banerjee和Duflo,2009)。例如,针对随机化偏差和试点偏差带来的挑战,班纳吉和迪弗洛及其合作者借助小规模补习的实地实验通过反复的设计、实验、再设计和再实验这样一个迭代过程扩展到大规模的政府主导的教育体系中。他们这个实地实验的最初想法来自一个印度非政府组织开发的一种教学方法,即教师根据学生的知识水平而不是按规定年龄等级的教学大纲教授基础语言和数学。班纳吉和迪弗洛等人通过实地实验验证了这个教学方法对落后的学生在常规课堂之外接受语言和数学补习的教育效果。为了评估随机化偏差带来的外部有效性问题,他们在农村环境中开展了一项新的实地实验,发现新评估的结果是积极的,但也有额外的问题,比如学习分布靠后的学生的接受率低(Banerjee等人,2010)。随后,他们将该补习教育嵌入了公办学校系统内,结果令人失望。调查结果表明,公办教师有能力提供补习教育,但大多数教师选择不这样做,这是试点偏差带来的外部有效性挑战。他们对实地实验进行修正,设计了教师主导模式下确保教师将补习教育视为一项核心教学任务,同时还制定了校内志愿者主导模式的补充方案,并且重新进行实地实验评估,这个项目表现出了积极的效果。经过长达数年的五项随机对照试验,他们设计了两个版本的补习教育方案,取得了大规模的成功。目前在印度13个州的10万多所学校实施了教师主导的模式,惠及近500万名儿童;有4000多所学校实施了由校内志愿者领导的模式,惠及20多万名儿童。

总而言之,他们对导致贫困的具体机制以及减轻贫困的有效干预措施的实地研究取得了丰硕的成果,使人们对哪些措施能够有效地解决贫困问题有了更清楚的认知。在他们的推动下,发展经济学关注的焦点逐渐从宏观层面的研究转向微观层面的研究,即从微观层面探讨贫穷国家如何改善健康、教育、工作条件、国内和国际政策以及市场条件等因素,从而为发展经济学研究奠定了坚实的微观基础。他们的研究说明,了解并最终减轻贫困的一个核心步骤是,找

出观察到的低效率的根源以及可以解决这些问题的政策;同时,他们对贫困问题解决方案的设计,也表明激励机制的重要性,而要设计有效的机制,就要找出处在贫困中的人们真正重视的东西是什么,据此才能设计出有效的政策来影响和改变他们的行为,进而改变我们的世界。

(三)新实证发展经济学得失评价

那么,如何评价他们的新实证发展经济学特别是对贫困问题具体研究的贡献呢?我们认为,应该将其放入发展经济学发展演变的脉络中,这样才能更好地把握其在发展经济学中的地位及其问题所在。虽然发展经济学作为一个独立的学科起于20世纪40年代中期,但是经济发展一直以来都是经济学家关注的重要话题。以威廉·配第和亚当·斯密等人为代表的古典经济学家主要关注劳动分工对经济增长的含义,认为劳动分工是经济增长的源泉,分工的程度依赖于市场的大小,市场的大小又取决于运输条件;劳动分工、专业化、货币的出现、投资与资本积累、商业扩张以及城市化等等各种现象之间与经济学发展存在着内在的联系(杨小凯,2003)。在古典经济学家看来,经济发展是一个劳动分工和市场演进相互作用的过程,在这个过程中,劳动分工的演进与影响交易成本的制度变迁相互影响,涉及劳动分工、专业化、市场网络、工业化、企业、贸易、技术进步、城市化等一系列广泛的相关议题,构成了丰富多彩的古典经济学发展思想。而现代发展经济学则对劳动分工与市场演进的相互作用的过程漠不关心,对经济发展过程中的许多相关议题视而不见,转而考虑诸如产业政策、投资政策、贸易保护政策等国家主导的发展战略,这实际上考虑的是给定资源条件下的最优配置问题。而面对国家主导的发展战略的失败,发展经济学家转而将目标聚焦于具体的贫困问题。

从对经济发展研究的视野及议题的广泛性来看,与古典经济学家的经济发展思想相比,现代发展经济学聚焦于给定资源的最优配置问题,而对劳动分工与市场研究相互作用的过程以及由此引发的包括制度变迁相关的种种现象视若无睹,实际上使得发展经济学的视野变得狭窄。面对现代发展经济学的失败,新实证发展经济学家不是将目光转向经济发展机制的研究,而是转而关注具体的贫困问题,可以说是发展经济学视野的进一步狭窄化。在以亚当·斯密为代表的古典经济学家看来,经济发展是劳动分工与市场演进相互作用的过程,贫困问题的解决是经济发展的题中应有之义。而新实证发展经济学家关注贫困问题,反而可能对贫困问题的根本解决产生有害的影响。

从我国的经济发展和扶贫实践来看,可以更清楚地看出新实证发展经济学的不足。改革开放以来,我国的经济通过改革开放改善了交易的条件,从而将广大的人民群众纳入到劳动分工中来,扩大市场网络,促进了专业化、企业、贸易、工业化等的发展,促进了技术的进步,而这个过程也涉及城镇化、城市化等相关的问题。从这个意义上讲,我国改革开放以来的经济发展可以说是古典经济发展思想的复兴,大部分的贫困问题正是在这个过程中得到解决的。反观三位诺贝尔经济学奖得主在发展中国家进行的实地实验,由于缺乏大环境的改善,贫困问题虽然有所缓解但是并没有得到根本的解决。以他们对提高教育质量的实地实验研究为例,发展中国家的教育问题与教育的回报率低有关,而教育回报率低则是由于分工合作水平不够高以及市场范围不够大。只有足够高的分工水平和足够大的市场,才能使得教育获得高额的回报,由此才能激励人们努力去获取高质量的教育。

总之,贫困的根本解决应该通过改善交易条件、降低交易成本将广大的贫困人口卷入到社会分工的大熔炉中来,在劳动分工与市场演进相互作用的过程中,通过专业化、工业化等方式去解决。当然,演进的过程必然要变革以往的生产方式和生活方式,这个过程不可避免地会带来新的问题,这些问题正是发展经济学家应该关注的议题,发展经济学应该回归和重视经济发展机制的研究。

五、结 语

自20世纪40年代中期发展经济学作为独立的学科创立以来,发展经济学的研究经历了盛极一时、发展停滞再到微观化的发展过程。在不同的发展阶段,学者们以不同的理论从不同的角度讨论了落后国家工业化、减少贫困人口等诸多问题。由于早期发展经济学家对发展中国家的复杂性、多样性缺乏足够的认识,名噪一时的“宏大的理论”并未解决发展中国家的贫穷问题,往往是旧的贫困问题还没有解决,新的贫困问题又出现。其实,贫困产生的原因极其复杂,正如阿马蒂亚·森所言,有许多关于贫困的事情是一目了然,但是其最终原因却是模糊不清的,是一个还远远没有定论的问题(Sen, 1981)。20世纪90年代以来,随着经济学研究从理论研究转向实证分析潮流的兴起,班纳吉和克雷默等也从理论研究转向实证分析。他们以实地实验为基础的新实证发展经济学,无疑是给发展经济学的发展和应用注入了新的生机。

一是促进有着自然科学规范的实地实验方法在发展经济学以及社会科学领域的应用。实地实验方法与工具变量法、回归断点设计、双重差分法、匹配等一起推动了计量经济学可信性革命的潮流,丰富了微观计量的分析方法,使得这些方法在发展经济学中得到广泛的应用。如前文所述,实证的关键是因果关系:自然科学先拿小白鼠做实验,然后推广到人身上还要做实验,社会科学也是一样,计量经济学要找偶然实验模仿自然科学,然后进行稳健性检验,而实地实验直接按照自然科学的规范做实验,然后在推广的时候也要考虑外部有效性问题。他们利用实验方法对健康、教育以及农村小额信贷等多方面的问题进行了研究。同时,其研究视角也扩展到社会、法律、文化、制度等非经济领域。从这些研究中可以看出实地实验方法有着其他方法不具备的优势,因此,经济学界应该把实地实验方法纳入因果关系检验的工具箱之中。

二是给减贫实践及公共政策制定提供了新的思路。他们的实地实验工作,对于发现贫困人口的行为动机以及找到有效的策略有很多贡献。由于实地实验的实验过程贴近真实世界,相对于传统的实证分析方法,能更为直接和便利地对变量之间的因果关系做出检验(罗俊等, 2015)。通过实地实验方法验证具体减贫政策的效果,实际上是提高了减贫措施的精准性,间接降低减贫的成本。这正如格尼茨和李斯特所指出的,政策设计者和企业常常会把实验想象成一件代价高昂的事情。但是,不进行必要的实验,才会付出更大的代价(尤里·格尼茨和约翰·李斯特, 2015)。实地实验的方法也可以应用于经济改革、企业经营、公共政策制定等领域。譬如,基于实地实验的研究设计,通过比较在相关公共政策影响下的人群和控制组人群的表现差异,来检验公共政策的效应,这样更能因地制宜地制定有效的管理方案,提高公共管理政策的效率。

但是,他们的研究也存在着诸多的争议。其中,最大的争议在于发展经济学理论与实证之间的关系问题——与发展经济学的实证研究相比,发展经济学的理论研究是否太少?^①。有西方学者直言不讳地指出,当前的发展前沿中理论研究太少(Mookerhjee, 2005)。毋庸置疑,尽管发展经济学研究微观化具有趋同态势,但其局限性也是显而易见的。

一是实地实验注重单一政策措施效果的检验,但缺乏理论的创新发展。客观地看,从古典发展经济学到新古典发展经济学,其研究的视野逐步变窄,而从新古典发展经济学从理论到实证再到新实证,视野就变得更窄。杨小凯(2003)认为,所谓的经济发展实际上应该看作是劳动分工与市场演进交互促进的过程,它涉及劳动分工、专业化、市场网络、企业、贸易、技术进步、城市化等一系列相关的问题。贫困发生与经济不发展是一枚硬币的两面,如果经济发展了,贫困在经济发展的过程中一定会得到减轻。由于实地实验过于关注细节,缺乏对发展中国家存在的经济政治问题和贫困发生的内在机制的探索,对发展经济学的理论创新可以说没有太多的

^①《比较》(第28辑)在“发展经济学专题”组织了五篇论文,讨论了发展经济学的方法论问题,相关的论文参见《比较》(第28辑),吴敬琏主编,中信出版社,2007年1月。

建树。贫困问题的根本解决需要将广大的贫困人口纳入到社会分工网络中来,而这个过程不可避免地会带来新的问题,这些问题正是发展经济学家应该关注的议题。

二是实地检验过于具体地盯住贫困问题的各个因素,缺乏整体性、系统性、制度性的研究。随着经济发展内涵的不断丰富,贫困的内涵也在不断地深化。贫困视角逐步转向经济、社会、政治、文化等多维视角,从微观的角度去解决贫困问题是远远不够的。譬如我国,改革开放40年来,我国展开了有组织、有计划的大规模扶贫工作,特别是党的十八大以来,脱贫攻坚工作力度之大、规模之广、影响之深前所未有。这是一个系统性、整体性的减贫工程。东部各个发达省市、社会各个部门的共同参与和彼此协调,使得7亿多的贫困人口脱离了国际贫困线标准。扶贫的巨大成就令世界瞩目,谱写了世界扶贫史上的绚丽篇章。它既体现了社会主义制度的优越性,也体现了国家“农转工”整体发展战略的成功。总之,经济发展与减贫是一个复杂的系统工程,过于关注单项的实验与设计,过于注重细节的研究,缺乏导致贫困的制度分析,难以建立解决贫困问题的长效制度保障。

主要参考文献

- [1]阿比吉特·班纳吉,埃斯特·迪弗洛著,景芳译. 贫穷的本质: 我们为什么摆脱不了贫穷[M]. 2版. 北京: 中信出版社, 2018.
- [2]杰弗里·萨克斯著,邹光译. 贫穷的终结: 我们时代的经济可能[M]. 上海: 上海人民出版社, 2007.
- [3]罗俊,汪丁丁,叶航,等. 走向真实世界的实验经济学——田野实验研究综述[J]. 经济学(季刊),2015, (3): 853-884.
- [4]威廉·伊斯特利著,姜世明译. 在增长的迷雾中求索[M]. 北京: 中信出版社, 2005.
- [5]杨小凯著,张定胜,张永生译. 发展经济学: 超边际与边际分析[M]. 北京: 社会科学文献出版社, 2003.
- [6]尤里·格尼茨,约翰·李斯特著,鲁冬旭译. 隐性动机: 日常生活中的经济学和人类行为背后的动机[M]. 北京: 中信出版社, 2015.
- [7]Banerjee A, Karlan D, Zinman J. Six randomized evaluations of microcredit: Introduction and further steps[J]. *American Economic Journal: Applied Economics*, 2015, 7(1): 1-21.
- [8]Banerjee A V, Duflo E. Growth theory through the lens of development economics[J]. *Handbook of Economic Growth*, 2005, 1: 473-552.
- [9]Banerjee A V, Duflo E. The experimental approach to development economics[J]. *Annual Review of Economics*, 2009, 1: 151-178.
- [10]Banerjee A V, Duflo E, Glennerster R, et al. Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives[J]. *BMJ*, 2010, 340: C2220.
- [11]Chattopadhyay R, Duflo E. Women as policy makers: Evidence from a randomized policy Experiment in India[J]. *Econometrica*, 2004, 72(5): 1409-1444.
- [12]Duflo E, Dupas P, Kremer M. School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools[J]. *Journal of Public Economics*, 2015, 123: 92-110.
- [13]Duflo E, Hanna R, Ryan S P. Incentives work: Getting teachers to come to school[J]. *American Economic Review*, 2012, 102(4): 1241-1278.
- [14]Duflo E, Kremer M, Robinson J. Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya[J]. *American Economic Review*, 2011, 101(6): 2350-2390.
- [15]Kremer M. Randomized evaluations of educational programs in developing countries: Some lessons[J]. *American Economic Review*, 2003, 93(2): 102-106.
- [16]Kremer M, Rao G, Schilbach F. Behavioral development economics[J]. *Handbook of Behavioral Economics: Applications and Foundations 1*, 2019, 2: 345-458.
- [17]Miguel E, Kremer M. Worms: Identifying impacts on education and health in the presence of treatment externalities[J]. *Econometrica*, 2004, 72(1): 159-217.

- [18]Mookerjee D. New directions in development economics: Theory or empirics? - Is there too little theory in development economics?[R] Boston University - Department of Economics - Working Papers Series WP2005-028, 2005.
- [19]Sen A. Issues in the measurement of poverty[A]. Ström S. Measurement in public choice[M]. London: Palgrave Macmillan, 1981.

The Field Experimental Approach to Causality and the Way of Poverty Alleviation in New Positive Development Economics: A Review of Main Contributions by 2019 Nobel Economics Laureates

Li Baoliang¹, Guo Qiyu²

(1. School of Economics and Finance, Huaqiao University, Quanzhou 362021, China;
2. School of Economics, Xiamen University, Xiamen 361005, China)

Summary: Three economists, Abhijit Banerjee, Esther Duflo and Michael Kremer, were awarded the Nobel Prize in Economics in 2019 by the Royal Swedish Academy of Sciences for their experimental approach to alleviating global poverty. Their field experimental approach to causality and the conclusions they arrived at are of great importance for researchers, policymakers and business operators. This paper will review their outstanding contributions to their experimental approach to global poverty alleviation. Four main conclusions are presented as below: First of all, based on the comparisons of the similarities and differences of empirical research on causality between natural science and social science, we show that field experiments have the same basic principles as natural science and other social sciences, and they are all faced with disputes about internal and external validity. Second, field experiment, as a tool of causality inference, has special advantages in economic empirical research, so that it should be included in the toolbox of researchers and gain the attention it deserves. Moreover, field experiments can be used as an effective tool for government economic policy-making and business management practice, which should be promoted. Third, their field experiments on poverty and the conclusions they arrived at are of important reference value, which should be affirmed. In particular, they remind us that in the process of mechanism design or policy-making, the design of details is very important and must be carefully considered. Last but not the least, from the perspective of the evolution of development economics since Adam Smith and from the practice of poverty alleviation in China, they regard poverty as the core topics of development economics research, which makes the field of vision of development economics research too narrow and may be harmful to the fundamental solution of poverty. The fundamental solution of poverty needs to bring the vast number of poor people into the social division of labor, and this process will inevitably bring new problems. These are the issues that development economists should pay attention to. In short, development economics should return to and pay attention to the study of economic development mechanisms.

Key words: Abhijit Banerjee; Esther Duflo; Michael Kremer; field experiment; casualty; poverty and poverty alleviation; development economics

(责任编辑: 宋澄宇)