

迈克尔·克雷默对发展经济学的贡献^{*}

——2019年度诺贝尔经济学奖得主学术贡献简介

张延 王琪

摘要:因发展了减轻全球贫困的实验方法,迈克尔·克雷默与阿比吉特·班纳吉、埃丝特·迪弗洛共同荣获2019年诺贝尔经济学奖。克雷默以经济学理论为指导,将随机对照实验引入减贫问题的研究,为大量发展政策提供了经验证据,并提出了众多缓解全球贫困的有效措施。他的研究方法也有助于重塑发展经济学的研究范式,为实验经济学乃至整个实证研究的发展做出了重要贡献。本文围绕克雷默在教育、健康、行为偏误等多个方面的实验研究,对其学术成就进行梳理和归纳,以理解其在贫困问题研究和研究方法论上做出的开创性贡献。

关键词:迈克尔·克雷默 随机对照实验 发展经济学 减贫 诺贝尔经济学奖

一、引言

迈克尔·克雷默(Michael Kremer)是美国著名发展经济学家,哈佛大学经济学教授。克雷默1964年出生于美国纽约,1985年以优异成绩毕业于哈佛大学并获得本科学位,1992年获得哈佛大学经济学博士学位。毕业后曾先后任教于麻省理工学院(1993—1999年)和哈佛大学(1999年至今)。克雷默因对发展问题的研究贡献,曾获麦克阿瑟奖,被世界经济论坛评为全球青年领袖,当选为美国艺术与科学学院(AAAS)院士,2006年被《科学美国人》杂志评选为年度50名研究人员之一。2010年起,克雷默担任美国国际开发署(USAID)发展创新风险投资(DIV)的创始科学主任,同时还是精准农业发展委员会的一员,美国国家经济研究局(NBER)、发展研究和经济分析社(BREAD)研究员。2019年,克雷默与阿比吉特·班纳吉(Abhijit Banerjee)、埃丝特·迪弗洛(Esther Duflo)共同获得了诺贝尔经济学奖。

当今时代,尽管世界经济发展取得了巨大进步,但全球仍有7亿多人生活在极度贫困中,每年有大约500万五岁以下儿童死于一些可以通过廉价治疗来预防或治愈的疾病,有一半的儿童仍然没有掌握基本的识字和计算能力就离开了学校(Committee for the Prize in Economic Sciences in Memory of Alfred Nobel,下文简称 the Prize in Economic Sciences, 2019)。克雷默的研究集中于发展中国家的教育、卫生、水利和农业,也曾帮助制定了疫苗的先期协议市场承诺(advance market commitment, AMC),以刺激疫苗研发的私人投资并推动发展中国家疾病预防疫苗的普及。他从微观角度切入全球贫困问题研究,通过引入随机对照实验的方法设计科学的实地实验,评估不同措施的影响效果,试图寻找有效的干预措施来帮助贫困人口改善生活。他的实验研究成果可被广泛用于发展中国家进行教育补贴、教师管理、疾病预防、医疗定价、技术推广等现实问题中。克雷默开创并推广的实验方法本身也对经济学多个分支具有重大意义,不仅重塑了发展经济学研究,也为整个实证经济学领域提供了更为丰富可信的经验证据和科学有效的研究方法。

^{*} 张延、王琪,北京大学经济学院,邮政编码:100871,电子邮箱:zhangyan0989@163.com, wangqi0727@pku.edu.cn。感谢匿名审稿人的修改建议,文责自负。

本文将介绍克雷默对经济学研究方法论的创新性意义,并通过梳理克雷默在实验中得到的经验证据来理解其对全球发展与减贫工作的重要贡献。

二、随机对照实验:开启经济学研究的新范式

克雷默将全球贫困问题分解为与个体或群组相关的更为精确的问题,如提高入学率、提高疫苗接种比例等,试图在改善穷人生活状态的微观视角上给出解决办法,其采用一系列巧妙设计的随机对照实验,并寻求与发展中国家部分机构的合作以在真实环境中评估减贫项目的效果。这种新的实验设计方法不仅在全球减贫问题上具有重大的实践意义,也重塑了发展经济学的研究范式。

(一)随机对照实验的早期发展

随机对照实验(randomized controlled trial, RCT)最初源于医药学中用于检测某种疗法或药物效果的控制性临床实验。尽管早期的对照实验通过一些方法控制了可能影响实验结果的部分客观因素,以尽量提取由不同干预造成的结果差异,但并未强调实验对象需随机分组,这极大地削弱了观测结果的可靠性。直至20世纪前叶,罗纳德·费希尔(Ronald Fisher, 1925)在进行农业和生物进化领域的实验研究时,为避免选择性偏误引入了随机化原则,并基于统计理论规范了科学实验的设计方法,由此形成了更为严格的控制实验法——随机对照实验。科学的实验设计使得观测结果被更为广泛接受,也推动了随机实验方法在各个领域的应用,包括医学、药学、农学、生物学等,并逐渐渗透至社会科学的研究中。

20世纪80年代之前,实证经济学主要依赖回溯性观测(调查)数据进行估计和预测,这在评估政策影响时遇到了诸多限制(Duflo, 2006)。政策评估中推测因果关系面临的基础难题是由于在任一给定的时间点,个体只能处于参与或未参与该项目其中的一种事实状态,而缺乏与之相对应的反事实结果,例如难以得知参与了某一项目的个体如果未参与该项目将如何表现,这导致研究者无法通过比较事实结果与反事实结果来直接评估在特定时间某一项目对个体的影响。一个解决方法是对比受项目影响的群体与由相似个体组成的不受该项目影响的群组,估计项目对一组个体的平均影响,而要使得结果避免选择性偏误等风险,需首先将具有相似特征的个体随机分配至实验组和对照组,并严格控制可能影响结果的其他因素。至此,在其他学科实验室中发展了两百多年的随机对照实验的思想被引入社会实验,并逐步发展为实验对象在日常环境中做出选择的实地实验。大量的实地实验提供了丰富的观测数据和政策支撑,实验内容涉及收入税、健康保险、福利改革、教育等众多问题。

值得注意的是,虽然随机对照实验具有随机性、对照性等诸多优势,其在实地实验中的应用仍会面临很多复杂的统计问题,不同于实验室中高度可控的环境,在真实生活背景中实施的社会实验通常更难以控制和捕捉差异。例如,个体可能拒绝参加实验或中途退出实验,其影响需在实验设计及实验结果中加以考虑(Duflo & Kremer, 2003)。与其他领域的随机对照实验的另一个关键性区别是,在实地实验中,干预结果反映的是个体的选择,即经济领域的实验者所探究的因果关系是通过人的行为建立的,因此要理解某一特定干预的影响结果,需理解产生这一结果的个体决策(the Prize in Economic Sciences, 2019)。例如,随机对照实验在医学中被广泛应用于探究疟疾、腹泻等致命疾病的治疗方法,但在相关医疗知识非常丰富的今天,仍有大量的低收入国家的儿童死于这些可以很容易避免的疾病。这导致研究者开始思考为什么这些儿童没有受到并不昂贵的预防或治疗,如何以有效的方式克服相关医疗服务不可得的障碍以降低低收入国家儿童死亡率,这些涉及个体行为的问题无法由医学实验回答。因此,以克雷默等(Miguel & Kremer, 2004; Kremer & Miguel, 2007; Kremer, Leino, Miguel & Zwane, 2011)为代表的发展经济学家,通过一系列特定的随机对照实验给出了答案。

(二)基于随机对照实验的实地实验经济学

评估干预影响的研究都试图回答这样一个反事实的问题:参与了某一项目的个体在没有这一项目影响下将会如何?没有参与某一项目的个体若在这一项目中将表现何种特征?而在任一给定时间,被观测的个体只存在参与或者未参与该项目中的一种状态,因此难以获得项目对个体影响的估

计(the Prize in Economic Sciences, 2019)。若要通过对照实验评估项目影响,克雷默和迪弗洛(Duflo & Kremer, 2003)认为,关键的一步是建立可信的对比。以实验组代表受项目影响的个体组,对照组代表不受项目影响的个体组,则可信的对比是指,若实验组未参与该项目,将与对照组具有相似的结果。在此基础上,对照组可以预测出实验组个体在没有该项目影响下将是何种特征,通过这样的设计可以给出上述反事实问题在群组层面的答案,即项目的平均影响。

但在现实中,克雷默和迪弗洛(Duflo & Kremer, 2003)研究发现,通常参与了某一项目的个体与未参与的个体本身就有所不同,有些项目也仅允许符合某类特质的人群参与(例如扶贫政策仅针对贫困人口实施),另外,是否参与该项目的决策在很多情况下也是自愿的。这导致两组人群的差异不仅反映了项目的影响,还包含了预先存在的差异,即选择性偏误。因此,参与人群与未参与人群不能形成可信的对比,实验组和对照组不能简单等同于参与者和未参与者。消除选择性偏误的一个方法是在潜在参与人群中随机选择实验组和对照组。这种方法使得两组人群平均而言没有显著差别,两者的不同可被归因为参与项目状况所致。

基于对照和随机这两个基础思想,克雷默及其合作者(Vermeersch & Kremer, 2005; Kremer, Moulin & Namunyu, 2003; Glewwe, Kremer & Moulin, 2009)设计了一系列的实地实验,并开展于多个发展中国家,通过与当地机构合作设计干预项目,在实施后评估干预效果。他们与当地具有在现实中实施项目的权力并对如何改善项目感兴趣的机构(如政府、非政府组织、私人企业等)合作,这保证了实验在真实环境中实施。研究者和机构均从这一合作中获益(Duflo, Glennerster & Kremer, 2006)。一方面,研究者基于经济理论帮助其合作者设计可解决实际问题的项目或政策,同时也可检验理论的现实适用性;另一方面,政府部门通过与研究者合作,可以推出多个改革实验措施用于随机选取的区域或群体,并选出较好的政策加以推广;非政府组织也希望检测新项目的效果或评估现有项目的作用;私人企业出于更好地理解其商业活动及服务客户、提高利润的目的,同样需要与研究者的合作进行随机实验。

克雷默及其合作者的创新不仅在于将随机对照实验引入发展经济学,而且还在实地实验的具体情境中提出了很多方便研究者在现实中实现随机化的方法。如克雷默等(Duflo, Glennerster & Kremer, 2006)所述,经典的临床实验中只需将实验样本随机分为实验组和对照组,其中前者接受某一新的项目干预。实地实验也可以沿用这一思想,但现实条件常使得随机分组不易实现。例如,存在资源、行政约束或需评估的项目是现存而非新引入的,因此研究者发展出超额认购(oversubscription)、分阶段随机引入(randomized order of phase-in)等多种方法。超额认购是指当资源或实施能力有限且对项目的需求超过供应时,在合格的候选人中随机选择获得该计划的人,这既是一种公平的分配资源的方法,也自然地引入了随机化。克雷默及其合作者(Angrist, Bettinger, Bloom, King & Kremer, 2002)在哥伦比亚进行的一项学校票券项目便采用此种方法决定票券的分配。另一种方式是将引入项目的顺序随机化,当存在金融或行政限制时,一些项目不得不以阶段性引入的方式进行,最公平的方式是随机决定引入顺序,这允许了参与项目和未参与项目的个体或群体同时存在且是随机选出的,而且这种分期引入的承诺可以促使被随机选为对照组的个体或群体保持与研究者的合作。因为对未来收益的期望为受试者提供了合作动机,克雷默等(Miguel & Kremer, 2004)进行驱虫实验时便分为三个阶段,随机选择学校并向其提供驱虫药治疗,以评估驱虫治疗对学生健康和学校出勤率的影响。

总之,克雷默等人将随机对照实验引入经济学研究用来检验经济理论预测的结论,并通过评估不同项目的因果影响提出明确的政策建议。这一实验方法也发现了许多现有模型未能预测到的机制和事实,并促进实施部门设计出更有效的政策或项目。

(三)对发展经济学的贡献

迪弗洛(Duflo, 2006)在总结实地实验在发展经济学中的地位时提到,最初的发展经济学常依赖于搜集原始数据来检验特定假设或研究某一特定问题,因而在很大程度上是由研究的问题决定所需

数据,而非由数据决定可以做的研究。但20世纪后期出现了众多高质量、大规模的数据集,如普查数据和不同领域的调查数据,丰富的数据来源和自然实验为发展经济学开拓了广阔的研究可能性,也促使更为专业的计量方法被引入,发展经济学采用实证证据的标准被大大提高,而基于原始观测数据的研究不再像之前容易被接受。在此背景下,克雷默等(Vermeersch & Kremer, 2005; Kremer, Moulin & Namunyu, 2003)在肯尼亚发起了一系列与教育相关的实地实验,在现实环境中引入随机对照的思想,通过对比随机分配至实验组和对照组的个体来评估干预影响。这种科学严谨的实地实验既传承了发展经济学从调查中搜集原始数据的方法,又重塑了原始数据调查的权威性和科学性,使得实地调查的结果得以被学术界认可,延续了发展经济学先提出问题后寻找相应数据来回答的传统,再次奠定了实地调查在发展经济学中的地位。一旦建立了这一方法的可行性,其在不同项目、国家、时期等情境下的多样性和可实践性便突显出来,相对于回溯性计量研究的优势也得以被承认。

在方法论上,随机对照实地实验最大的优势是可以构造任何调查数据或自然实验均无法观测的反事实结果,这得益于其灵活的实验设计和干预选择(Duflo, Glennerster & Kremer, 2006)。观测研究只能评估现实世界中已经存在的事实结果,而实地实验可以通过实施干预创造出自然中无法观测到的事件,从而研究这一干预的影响。这使得研究者在政策尚未实施时即可检验其是否如先验知识或理论预测的那样有效,也可以在检验理论预测时通过设计不同的干预引入随机变化。

在所研究的问题上,随机对照实验在发展经济学理论与政策适用性之间建立起直接桥梁,将贫困背后的主要机制与对于不同政策干预的行为反应融合起来探讨(Duflo, 2006)。实验中干预的设计是由发展经济理论指导的,收集的观测数据是基于微观个体选择的,因此实验结果可以指明理论所发现的机制及政策方向在现实中的有效力,定量地测算某一因果关系,为理论预测提供更有说服力的实证检验。

(四)相对非实验方法的优势

为控制选择性偏误实现反事实推断,基于非实验数据发展出了倾向得分匹配、双重差分法、断点回归法等政策评估方法,但克雷默(Duflo & Kremer, 2003)指出,这些非实验方法均有各自的局限性。倾向得分匹配法是根据可观测的个体特征预测其参与概率,以此为依据为受干预影响的实验组匹配参与倾向相似的对照组,但这一方法依赖于对受干预和未受干预两个群组自身差异的准确识别,这导致当基于某一难以观测的特征实施干预时,估计可能存在偏误。双重差分法通过比较受政策影响的某一区域和未实行该政策的区域的增长差异来探究该政策的影响,但结果的可靠性所依据的假设前提是:若没有该政策,两地发展趋势相同。而研究者无法检验这一假设,另外还要保证同一阶段两地没有实施其他可能影响结果的政策,以及未受到任何外来持续性冲击,这在现实中均难以实现(Glewwe, Kremer, Moulin & Zitzewitz, 2004)。断点回归所比较的是某一临界值附近参与项目和未参与项目的个体,认为两类群体其他特征相似,结果的差异仅来源于是否参与了项目,但这只适用于可产生此类断点(临界值)的项目,而在具体实施中断点常常是不存在的,尤其在政策实施不严格的发展中国家。而即便存在断点,这一方法也很有可能面临错误估计标准误差的风险(Bertrand, Duflo & Mullainathan, 2004)。

随机对照实验则弥补了非实验的干预效应研究方法的缺陷,使得实证分析在更为广阔的领域和问题上具有可信的结果(Duflo, Glennerster & Kremer, 2006)。实验中干预的实施可供研究者设计并操作,随机化的处理可以控制选择性偏误,实验组和对照组的差异也可以被精准记录。整个干预影响及观测过程均可保证科学透明,研究者可以直接获得高质量的对比数据,而无需对样本自身或计量方法做出过高的要求。克雷默等(Glewwe, Kremer, Moulin & Zitzewitz, 2004; Glewwe, Kremer & Moulin, 2009)比较非实验方法和随机对照实验的评估结果发现,两者会有很大差异,其中非实验方法更容易面临选择性偏误、遗漏变量等识别问题。随机对照实验的引入发展出了一套更为可信的实验设计过程,极大降低了实证中常面临的选择性偏误的风险,使得实验成为经济学研究

在检验理论和评估政策时的一种新的强有力的工具。

另外,克雷默等人创建了一套可迭代的实验研究过程(the Prize in Economic Sciences, 2019),所谓可迭代即研究者可以在相同环境下使用相同的结果变量和相同的测量技术进行后续研究,可以在已有的实验基础上研究引申问题。实证研究中经常会遇到需要进行后续研究的问题,尤其是基于观测数据的研究通常难以作为其他研究的基础被进一步推进,因为初始政策改变时所产生的数据有限,研究引申问题时需要的数据可能并不在这一观测集中,也无法保证其他条件相同来观测在初始政策基础上施加新的干预将产生何种新的影响。而实验具有可以控制干预发生的特点,使得研究者可以进行多步骤的序贯实验(sequential experiments),其中每个新的步骤都可以利用较早步骤中的结果,克雷默等(Duflo, Kremer & Robinson, 2008, 2011)关于肯尼亚农民化肥使用情况的发现便基于一系列长期序贯的实地实验。

总之,克雷默将随机对照实验引入经济学研究,形成了一种新的实地研究方法,以可信的方法实施大规模的实验,大大增强了经济学者研究现实问题的能力。

三、随机对照实验在减贫研究中的应用

克雷默在现实世界中实施了一系列的经济学实验,内容涉及教育、健康、理解贫困人口的行为偏差等。实验由真正生活于其中的个体做出关乎自身利益的决定,因此可以观测到更直接、更贴近现实的行为结果,大规模样本的复杂性也使得研究者可以理解个体间的相互影响和作用,这在减贫实践和理论研究中具有重要意义。

(一)寻找教育减贫的有效措施

自20世纪80年代后期起,宏观经济研究开始强调人力资本在增长和发展中的作用,相关的理论和实证研究常用教育程度近似代表人力资本。在考虑了人力资本的作用而进行增长核算时,重要的一步是衡量人力资本回报率,实证工作中常用明瑟收入函数(Mincer earnings function)估计的教育收益率代替。但这一方法面临三个问题(the Prize in Economic Sciences, 2019):首先,教育水平在不同部门间的差异是个体选择的结果,这导致基于截面数据来估计教育回报率通常是有偏的。其次,回报是根据受教育年限衡量的,但就人力资本获取而言,回报在不同时间和地点上并不相等。最后,这一方法对政策改进的作用非常微弱。90年代中期,克雷默及其合作者(Vermeersch & Kremer, 2005; Kremer, Moulin & Namunyu, 2003; Angrist, Bettinger, Bloom, King & Kremer, 2002)发起了发展经济学的变革,通过随机化的实地实验探讨了与教育相关的更为精准细微的问题,也正由于教育对于个体实现收入突破、国家实现快速增长的重要作用,克雷默等人在现实环境中评估了多种促进贫困地区适龄儿童接受教育的措施。

1. 提高入学率。发展中国家需要解决的一个基本问题是以尽可能低的项目成本实现提高入学率的目标。很多贫穷国家的适龄儿童因为要工作挣钱而不去上学,或者由于无法支付学费而辍学,因此提高入学率最简单的方式是降低上学成本。舒尔茨(Schultz, 2004)发现基于孩子是否上学向家庭提供现金补贴可显著提高入学率,但这一方法的筛选条件过于表面,在实践中常常难以保证严格将补贴发放给上学家庭,需要探讨其他可有效提高入学率的办法。

(1)提供免费早餐。学校向学生提供免费食物可以保证这一福利仅提供给在学者,克雷默等(Vermeersch & Kremer, 2005)与一个非洲的非政府组织 ICS(International Christelijk Steunfonds Africa)合作在肯尼亚发起了这一研究项目,他们在肯尼亚选取了50个幼儿园并随机选取其中的25个作为实验组提供免费早餐,另外25个幼儿园作为对照组。实验结果表明,实验组入学率比未提供免费早餐的对照组高30%,测试分数比对照组高0.4个标准差。

(2)降低学习成本。在很多国家,教育对家庭来说有很多私人成本,包括交学费以及购买校服、教材等,通常这对贫困家庭来说是一笔很大的开支。为探究入学率对教育成本的敏感度,克雷默及其合作者(Kremer, Moulin & Namunyu, 2003)选取了14个肯尼亚的学校,随机选取其中一半提供

校服、教材及教室建设费用以减轻家庭负担,另外一半不接受处理作为对照组,研究结果发现,这一项目使得实验组学校辍学率有显著下降。另外,这一教育成本的下降也使得附近学校的学生转学至实验组学校,导致其班级规模增加了50%。这表明学生和家長愿意放弃较小的班级规模来换取免费的校服、教材和教室建设。研究者进一步发现,同时减免额外费用和增加班级规模对学生成绩基本没有影响。这一实验指出肯尼亚政府应该放弃部分为减小班级规模所做的努力而增加对教材、教室建设等教育成本的资金支持,这一政府预算的重新配置可以大大减少家長的教育花费并提高参学率。降低学习成本的另一个直接方式是向学生提供金融支持,克雷默等(Angrist, Bettinger, Bloom, King & Kremer, 2002)在哥伦比亚进行了一项大规模的实地实验,向随机抽取的私立学校的小学生提供可部分覆盖其中学成本的票券,若取得了满意的成绩则票券可持续提供,通过比较获得票券和未获得票券的学生可以发现,有票券激励的学生出勤率和成绩均有所提高,而且出去工作、早婚或同居的概率更低。

(3)健康。健康问题也是影响入学率的一个重要因素,例如肠道蠕虫在学龄儿童间感染率很高,克雷默等(Miguel & Kremer, 2004)评估了一个基于学校提供驱虫药的项目对参学率的影响,75所学校以随机的顺序逐步加入该项目,研究结果发现,参与项目的学校旷课率比对照组学校旷课率低25%。另外,由于减少了疾病的传播,这一项目还有正外溢效应,考虑到这一影响,该项目将实验组中每人的受教育年限增加了0.15年。

2. 改善学校表现。克雷默等(Glewwe, Kremer & Moulin, 2009; Glewwe, Kremer, Moulin & Zitzewitz, 2004; Glewwe, Ilias & Kremer, 2010)的研究还关注如何改善学生在学校的表现,例如增加学校投入以向学生提供学习的配套设施是否有利于提高成绩,如何激励教师改善教学效果,如何鼓励学生更重视学习等。这方面的实验证明此前一些非实验研究的结论并不可靠,也指出了一些理论建议在实践中可能遇到的问题。

(1)增加学校投入。很多非实验的分析表明,给初始条件差的学校提供额外的教材可以改善学习效果,但克雷默等(Glewwe, Kremer & Moulin, 2009)在肯尼亚的随机对照实验中发现事实并非如此。研究者从100个肯尼亚农村小学中随机选取了25个向他们提供教材,发现额外的教材提供并没有导致实验组和对照组的平均测试成绩产生显著差别。虽然在项目实施之前本身就领先的学生成绩提高了,但之前成绩落后的学生并未受益,这一措施对于降低留级率和辍学率的作用也很微弱。研究者认为这可能与肯尼亚的具体国情相关,因为肯尼亚教学所用的官方语言是大多数小学生的第三语言,即便有了额外教材,成绩落后的学生也难以理解其中内容,而且课程内容对于缺乏高水平老师的农村学生过于难懂。进一步地,研究者基于学校的回顾性数据进行了截面分析,通过比较实验方法和非实验方法的差异发现,非实验估计错误地得出了教材对提高测试成绩具有显著正影响的结论。类似地,向学生提供配套挂图也未能有效帮助学生提高学习成绩(Glewwe, Kremer, Moulin & Zitzewitz, 2004),这一措施并没有表现出像非实验数据预测的正向促进作用。

官方教育内容与大多数学生需求不匹配导致单纯增加学校投入并不能产生预期的效果,而只是加剧了学生间成绩的分化,这种教育资源向优质学生集中的系统在发展中国很常见。克雷默等(Glewwe, Kremer & Moulin, 2009)认为,其原因是很多如肯尼亚一样的发展中国家出于某种政治目标常采用一种中心化、统一的国家课程和教育系统,专业化、针对性的教育(如职业学校)发展不足,难以对不同的学生进行定向培养。另一方面,教育快速扩张的过程中,学生间的异质性增强,学校课程体系未进行相应调整,随着入学率提升,尽管肯尼亚确实进行了一些课程改革,但在精英阶层掌握政治权力的系统下,大部分课程仍是针对少部分精英学生的,教师对优质学生的偏好也加剧了这一扭曲。克雷默的研究指出了发展中国家需关注教学内容的结构性调整,而不仅仅是教学投入的增加,也强调了对贫困地区的发展援助要与其需求相匹配。

(2)改革教育机制。如果增加学校投入对学习效果的影響均不尽人意,那么是否可以通过提升教育工作者的积极性来帮助贫困学生实现教育脱贫呢?事实上,一些发展中国家的教师并没有很好

地履行职责,他们的平均缺勤率高达 19%(Chaudhury, Hammer, Kremer, Muralidharan & Rogers, 2006),而且很大比例的教师并没有从事教学活动,克雷默试图寻求有效的办法激励教师更好地投入教学工作,让贫困地区的学生接受更好的教育。

一种方式是基于学生成绩向教师提供奖励,克雷默等(Glewwe, Ilias & Kremer, 2010)在肯尼亚进行了这一项目的实地实验,接受项目的学校学生成绩在项目期间确实提高了,但这种优势在项目结束后就消失了,分析原因时发现,接受项目学校的教师并没有增加自身出勤率,也没有改变教学方法,而只是增加了提升短期测试分数的备考时间,这表明基于成绩来对教师进行激励只会让教师“为了考试而教学”(teaching to the test),而并没有致力于提升学生的长期竞争力。

为寻求更加长效的激励方式,克雷默等(Duflo, Dupas & Kremer, 2015)考虑了一种短期合同制聘用制度和地方(主要指父母)监管制度。2003年肯尼亚推行初等教育免费时,一年级入学人数剧增,为减小一年级的班级规模,学校希望雇用额外的短期合同制教师教授新增加的班级。不同于原来的公务员系统下稳定的教师职位,合同制教师雇佣合约按年签订且基于当年表现可以续签,他们的工资也相对较低。在理论上,这种条件性雇佣合约依赖于动态激励(Holmström, 1982),若表现得好将被持续雇用,而表现不好将面临合约终止,因此合约教师应该比公务员系统的教师更愿意付出努力以留住工作。另外,由于短期合约的教师工资一般较低,很多发展中国家为了节约成本也更倾向于雇用合约制教师。理解短期合同制和公务员制教师对学生成绩的影响对理论检验和现实依据均有重要意义。

克雷默等(Duflo, Dupas & Kremer, 2015)设计了额外教师项目(Extra Teacher Program, ETP),随机选取了 70 所学校实施基于表现的短期合同制,学生被随机分配为留在原先班级或进入新的班级。在参与了 ETP 的学校中,研究者又实施了以学校为基础的管理项目(School-Based Management Program, SBMP),实验组引入了新的管理方式——赋予家长参与监管的权利,如面试老师和决定录用权、监测教师努力程度和表现、监管学校管理和审查某一合同制老师是否可以续签等。对照组则未采用这一形式,而是由现有的公务员老师负责录用和监管合同制老师,这一项目旨在考察将关于学校运行的决定权转移至地方学术委员会时是否应该增加家长的参与。ETP 和 SBMP 是一个多处理实验,可同时分析不同聘用制度和监管方式的影响。实施 ETP 的结果发现,留在原来班级继续由公务员老师教授的学生成绩没有显著增加,尽管班级规模平均由 82 人下降至 44 人,而合同制教师所带班级的学生成绩有显著提升,这类老师的出勤率也相对较高。实施 SBMP 的结果发现,招聘权下放给学校原有的公务员系统教师后,被录用的新老师约有 1/3 是原有老师的亲戚,同时原有老师的出勤率也有所下降,但让家长参与监管可以大大规范这一过程,并正向作用于学生成绩。克雷默等的研究验证了短期条件性合约的动态收益,也强调了监管的重要性。

(3)优化师生匹配。发展中国家常常面临教育资源稀缺的问题,教师不能分配足够的精力关注到每一个学生,为其提供合适的引导。简单的方法是通过雇用更多的老师和分班来改善师生比,增加每位老师在每个学生身上的可用时间。但正如上文 ETP 实验所示,单纯的减小班级规模对学生学习效果的提升作用非常有限,于是研究者考虑了另一种帮助老师因材施教的方式:根据学生初始成绩分班。但由于同龄人效应的存在,这一方式是存在争议的——若学生可以从优秀的同龄人身上获益,那么基于学生能力分班虽然可以有益于优秀学生,但对落后学生是不利的,这会加大学生间成绩的不平等。但另一方面,这可以使得老师更精准地基于学生需求给予相应的指导。克雷默等(Duflo, Dupas & Kremer, 2011)建立了一个涵盖这两种效应的模型,学生水平可以直接对同学产生外溢效应,同时也会间接影响教师的努力和对教学难度的选择,教师的选择取决于学生测试成绩的分布,以及教师所得回报是测试分数的线性函数、凹函数还是凸函数;学生的收益则取决于自身水平与教师教学内容的距离,距离越远获益越少,如果这一距离过大,学生将完全不会受益于教师的教学。基于这一模型,研究者用在肯尼亚的实验数据进行检验发现,根据初始成绩分班的方式可使得所有学生均从中获益。

具体来说,该项目实验样本为2005年肯尼亚121所小学的一年级班级,随机选取出60所学校作为实验组将学生基于初始能力分班,另外61所学校构成对照组将学生随机分班。18个月后,实验组的学生比对照组学生成绩高0.18个标准差,无论高能力学生还是低能力学生的成绩均高于对照组的相应学生。进一步地,作者还发现在高能力班级排名靠后的学生与低能力班级排名靠前的学生获益相同(排名均指基于基期的初始成绩分布下的排序),由此可以推断教师基于班级学生水平调整了教学内容,且教师的收益函数是班上学生最终成绩分布的凸函数,即教师设计教学时更关注于成绩优秀的学生。这使得初始成绩在中位数附近但低于中位数的学生被分到低水平班级后,尽管同龄人效应不利于他们成绩提升,但教师对他们关注的增加弥补了这一效应的缺失,初始成绩在中位数附近但更高学生在高水平班级相对缺少教师关注但受益于同龄人效应。另外,在随机分班方式下,若学生被随机分配到水平更高的班级,这一外生的同龄人初始水平提升对于顶尖学生有非常强的正效应,因为同龄人效应增强和教师指导更贴切;对于中游学生作用不明确,因为两种作用方向相反;对于低端学生有微弱的正效应,因为他们距离教师的指导目标过远,再增加错配程度的负效应较低,但他们可以从同龄人中获益。

3. 主要结论。在过去的20年,克雷默在发展中国家实施了大量的有关教育问题的随机对照实验,得到了丰富的实证结果:在低收入国家,学生对教育成本很敏感,减少自付费用、提供优秀学生奖学金和有条件的现金转移都可以促使贫困学生接受更高教育,解决儿童健康问题也可以经济高效地增加入学率。但是,如何改善学生在学校的表现是一个更复杂的问题,随机对照实验表明增加相同的教学输入(例如购买更多的教科书或提供配套挂图)对提升学生成绩的作用很小,减小班级规模、向教师提供基于教学成绩的奖励也收效甚微。相比之下,如克雷默等(Kremer, 2003; Kremer, Brannen & Glennerster, 2013)所述,对教师激励措施和问责制的改革(例如在当地聘用短期合同的教师)以及将教学与学生的学习水平相匹配的教学改革在促进学习方面是非常有效的。

(二)提高贫困人口的健康水平

当今社会医学的发展极大改善了人类的健康水平,即使发展中国家的居民也可以较为普遍地接受医学治疗和疾病预防,但低收入国家五岁以前儿童的死亡风险仍比高收入国家高15倍,一些低成本的疾病预防手段在发展中国家仍未全面普及(the Prize in Economic Sciences, 2019)。

克雷默等(Miguel & Kremer, 2004; Kremer & Miguel, 2007)结合理论与实验证据,从市场非有效的角度探讨了发展中国家公共健康投入不足的原因。健康是构成人力资本的另一个重要部分,理性人投资于健康时会比较其边际成本和边际收益,若前者小于后者则会增加投资,直至边际成本等于边际收益。但当某项投资的私人收益不能完全反映其社会价值时,个体选择便会偏离社会最优,例如对于一些具有正外部性或者公共物品性质的健康投入,其私人收益小于社会收益,这会导致个体自发进行的健康投入低于社会最优水平。

1. 驱虫治疗。肠道蠕虫感染了全球四分之一的人口,尤其易影响发展中国家的学龄儿童(Miguel & Kremer, 2004),公共健康和流行病学研究均试图寻找有效的方法治疗和预防这一疾病,但相关的实证研究却忽视了一个很重要的方面——传染疾病治疗的外溢效应。克雷默等(Miguel & Kremer, 2004)研究指出,在个体水平上随机分配药物治疗的研究可能会低估治疗的益处,因为忽视了减少疾病传播这一外部收益,从而低估了实验组接受治疗的收益。他们转为在学校层面分配药物并评估影响,估算驱虫药治疗项目的直接效应和外部性。

该项目中,驱虫药被分阶段随机引入以学校为基础的大规模治疗中,这一并非针对个人的治疗可以观测到项目的整体效应。75个小学的分阶段治疗顺序由一个列表确定,该列表先后按地理位置和字母顺序将所有学校分为三组。第一组于1998年和1999年都免费接受驱虫药治疗;第二组于1999年、第三组于2001年免费接受驱虫药治疗。因此,1998年第一组为实验组,而第二组、第三组为对照组;1999年第一组、第二组为实验组,而第三组为对照组。这是一个阶段性引入的随机对照实验。实验结果发现,驱虫治疗不仅有益于接受治疗的学校儿童,也极大改善了附近学校中未接受

治疗儿童的健康和学校参与程度,这证明了驱虫治疗确实存在很大的外部性,需要公共部门给予充分的补贴。

克雷默等(Kremer & Miguel, 2007)进一步考虑了这种健康产品的持续性提供问题,这也是发展经济学中的一项重要议题:如何将对外界援助的依赖转化为当地自发的可持续性行为。在上述免费驱虫药治疗项目中,他们试图寻找出可以代替补贴的可持续性抗虫措施,包括分担费用、健康教育、口头承诺(提前询问人们是否计划使用抗虫药物,这利用了社会心理学的发现,即个人会力求在陈述和行动中保持一致)和社会学习(若小范围补贴可以提高个体对新技术的接纳,其他人可以通过与这部分群体的社会联系进行学习并模仿采用),但均未能成功:驱虫药价格由零变为很小的正价格导致接受药物治疗率由75%下降至18%,需求对于价格是否为正非常敏感,但对具体为多少并不敏感,因此使人分摊很低的费用并不会产生预想的结果;健康教育和口头承诺均对人们的驱虫行为没有影响;增强与接受补贴群体的社会联系甚至降低了驱虫药的服用概率。正如他们所指出的,这些措施不尽人意很大程度上是由于存在外部性,即驱虫药的私人收益低于其花费,但他们未能将外部性从价格中剥离出来。

2. 水源保护治理。克雷默及其合作者(Kremer, Leino, Miguel & Zwane, 2011)基于水利基础设施研究了公共物品的供给问题,他们在肯尼亚实施的随机对照实验衡量了一项致力于泉水保护来改善水源质量的投資对当地居民健康的影响,估计了家庭对泉水保护的估值,模拟了不同水利产权制度的福利效应,包括共同产权、永久私人产权等。

该项目设计的理论基础来自公共物品的市场提供可能是非有效的。健康水源属于自然垄断产品,这种产品消费的竞争性很弱同时排他性又较低,水源保护基础设施的初始建设需大量固定成本投入,但一旦建成后,为额外的消费者提供服务的边际成本很低,即使这一服务是可排他的。如果家庭对水源保护的估值不同而提供者又不能实施完美的价格歧视,就会产生静态的低效率。但通过调整价格可以减少无谓损失,提高社会福利。

项目实施于肯尼亚西部的农村地区,当地很多家庭因为从天然泉水中采集食用水而面临严重的水源健康问题:尽管大多数泉水位于私人田地中属于私人所有,但根据肯尼亚的习俗和法律要求,泉水通常是免费提供的,因此泉水提供者没有动机去改善水源质量,这些水很容易被人类或其他动物的粪便污染。研究者基于此对实验组的泉水实施保护干预,包括基础设施建设、建立栅栏和排水系统、组织维护委员会等初始措施,以及之后的日常维护,如设备修补、清理等。研究者从当地符合条件的天然泉水中随机选取了200个接受实验,以阶段性引入的方式随机决定每个泉水接受实验处理的时间,每轮引入后接受处理和未接受处理的泉水形成实验组和对照组,直到最后所有泉水均实行了保护项目。实验结果显示,该水源保护项目将天然泉水中的大肠杆菌污染降低了66%,儿童腹泻率降低了约1/4。

此外,克雷默及其合作者(Kremer, Leino, Miguel & Zwane, 2011)还用实验数据估计了家庭对泉水保护的支付意愿。考虑到家庭在水源质量和取水地距离之间进行权衡来选择取水地,研究者采用离散选择模型估计出水源质量提升对家庭选择取水地的影响,以此测量家庭对更高质量的水源的偏好。进一步地,以这一需求系统为基础,研究者比较了不同产权制度下的水利基础设施投资和社会福利水平,以探究低收入国家水资源保护的最佳治理方式。研究结果表明,尽管家庭会增加受保护泉水的使用,但家庭的支付意愿并没有预测的那么高,其水源健康需求对收入是高度敏感的,因此,在低收入水平阶段,公共产权制度优于私有产权制度。但是当收入水平较高时,对于健康水源的需求会提升,私有产权可以刺激出足够的投资,会比公共产权更有效。

3. 主要结论。基于科学的实地实验,克雷默等人关注到困扰发展中国家的居民健康问题:一方面,精确测度了提升居民健康的投入如提供驱虫药、引入水源保护技术的作用,具有重要的现实意义;另一方面,为很多经济学理论预测提供了实证证据,如外部性、同龄人效应等。克雷默将基于理论设计的项目应用于实践中,实践结果反过来又为理论提供了检验。这一融合了经济学理论与现实

经验的证据表明, 流行病药物和公共物品的提供需要更多的政府支持, 发展中国家需根据居民收入水平来设置相应的产权制度。

(三) 理解贫困人口的行为偏误

行为经济学家发现, 在真实世界的很多情况下, 个体会偏离经济学模型预测的最优行为, 发展经济学也在发展中国家发现了越来越多类似的现象。例如, 非洲的很多小农场主并不愿意采用相对较为简单的现代技术(例如肥料), 尽管这些技术在农业实验中已被证明可以产生很高的回报(Duflo, Kremer & Robinson, 2008)。类似地, 一些发展中国家的企业组织也常会偏离利润最大化的决策。克雷默等(Kremer, Lee, Robinson & Rostapshova, 2013)研究发现, 肯尼亚农村的很多店主在进行存货管理时常会放弃一些回报很高的存货投资, 也未能考虑到平衡不同存货商品的边际收益。这使得发展经济学家思考在信息不完全的环境下, 即使个体是“新古典主义者”(具有无限理性、前瞻性和内部一致性者), 贫困是否会对其决策过程施加约束, 从而影响个体行为? 有限注意力、风险厌恶、自控力问题等行为偏误是否会影响减贫政策的效果?

克雷默等(Duflo, Kremer & Robinson, 2011)的工作是融合行为经济学和发展经济学的一个重要基石(the Prize in Economic Sciences, 2019)。为了探究为何可产生高回报的肥料在贫困地区未能普及, 他们于2000年开始在肯尼亚西部进行了长期的序列实地实验, 所发现的第一个原因是, 只有在正确使用的前提下, 肥料才是有利的, 但这对农民并不容易, 因此他们难以从中获利, 这导致当地只有大约20%的农民采用了这项技术(Duflo, Kremer & Robinson, 2008)。在后续实验中, 他们考虑了农民采用新技术时的拖延心理和时间不一致性问题, 这种结合了心理学和经济学对个体行为偏误的探讨极大地影响了后来关于贫困人口决策过程的研究(Duflo, Kremer & Robinson, 2011)。

在克雷默等(Duflo, Kremer & Robinson, 2011)建立的模型中, 农民具有一种现时偏误(present biases), 即会过度贴现未来的效用, 并低估自己在将来也不够耐心的风险。由于购买化肥需要成本, 尽管这一成本很小, 只要农民贴现率很高, 即使计划使用化肥的人也会推迟到截止日期前再买, 但到了最后阶段时, 他们将再次不够耐心, 然后放弃投资于化肥。这一模型对于有关化肥补贴的争论具有重要意义。当认为农民使用某一投入过少时, 农业专家常会主张大力补贴该项投入来鼓励农民采用。但经济学家历来对此表示怀疑, 认为这种过度干预可能会使得投入偏离最优水平而造成负面影响。克雷默等的模型表明, 尽管高额补贴可以帮助有现时偏误的农民克服购买化肥的直接效用成本, 但也会使得没有现时偏误的农民过度使用化肥。因此, 他们提出了一种更好的政策设计: 在农民收获后提供一种较小的、有时间限制的补贴。他们认为这种补贴可以减轻现时偏误, 同时对没有时间不一致性偏好的农民的行为仅有较小的扭曲作用。

基于这一模型, 研究者在项目中比较了以上两种政策干预: 一种是直接提供相对较大的化肥补贴, 在实验中具体表现为提供50%的打折并免配送费; 另一种是提供相对较小的化肥购买限时优惠, 如在居民刚好有钱的收获时期提供购买化肥免配送费优惠。实验结果发现, 第二种方式可以大大增加化肥的使用。这一结果与理论模型是一致的, 证明了现时偏误是贫困人口中技术采纳率低的重要因素, 而在个体有钱时, 提供限时折扣是一项有效的政策工具。

这一研究的里程碑作用不仅在于它是行为经济学与发展经济学的融合, 还开创性地发展了实验经济学的可迭代性。他们的研究建立在一系列序贯实验基础上, 即新一轮的实验可以回答之前实验的结果, 这种基于实验迭代学习的过程需要保持环境和人口一定, 通常仅可在实验室中发生(the Prize in Economic Sciences, 2019)。但克雷默等人的设计将实验室经济搬到了实地中, 使研究者可以观察真实世界中的决策, 为后续研究个体认知和心理决策开辟了新的道路。

四、实验结果推广面临的问题

克雷默等发展的实验方法的核心是建立因果关系, 一般认为随机对照实验在内部有效性(internal validity)问题上是可信的, 即可适用于实验中特定的干预、样本人群、时间和设计, 但同样的因果

推断可否推广至其他人群、时间和环境,即随机实验是否存在外部有效性(external validity)问题,需要进一步探讨(Duflo, Glennerster & Kremer, 2006)。本部分将指出可能影响实验结果推广的因素,以及克雷默及后续研究者在克服这一问题上的努力。

(一)均衡效应

由于随机对照实验比较的是在给定区域内实验组和对照组的区别,只能衡量某一项目对接受者的部分效应或当地效应,而未能包含一般均衡效应,这可能导致扩大项目规模时忽略一些重要的影响,例如价格和工资,也可能会影响项目的整体效果甚至会产生意想不到的分配结果。

对于一些可以预测的重要均衡效应,可以在实验设计时加以考虑。一种方式是在市场而非个体层面进行随机化,并关注市场层面的结果。另一种解决办法是可以进行两阶段的随机化实验,除了在实验市场内随机化实验,还要在市场层面施加随机性控制,迪弗洛等(Crépon, Duflo, Gurgand, Rathelot & Zamora, 2013)在研究就业援助的直接影响和间接影响时,第一步对参加实验的235个劳动力市场随机给定求职者比例,第二步在每个劳动力市场中,随机选取求职者接受就业援助的处理。这保证了不仅接受干预影响的个体是随机选定的,而且接受实验区域中求职者的比例也是随机的。

均衡效应也可能发生于更高的层面,比如国家或世界,此时难以进行随机化处理,但实验方法也可以估计一些关键的行为参数,用于分析更大规模下的均衡结果(the Prize in Economic Sciences, 2019)。

(二)溢出效应

溢出效应也可能在小范围实地实验中被忽视。所谓溢出效应是指研究者所关注的干预可能不仅对实验对象产生影响,还会对实验之外的群体具有外生效应。例如,上文提到的克雷默等(Angrist, Bettinger, Bloom, King & Kremer, 2002)在哥伦比亚进行的给私立学校学生提供票券激励的项目,可能通过增加私立学校的竞争力促使公立学校表现得更好,或者吸引重视教育、自身能力强的学生由公立学校转学至私立学校而不利于公立学校发展。忽视外部影响可能会高估或低估项目的投资收益,影响其定价和有效提供,也可能导致将项目由小范围扩展至更大范围时错误预估其效果。

与均衡效应类似,可以通过在更高群体水平上进行随机化实验来内化外部性,例如克雷默等(Miguel & Kremer, 2004)考虑了驱虫药治疗的外溢性,将实验设置于学校层面而非个人层面,将个体之间的外部性内生。另外,考虑到实验组学校接受治疗可能影响附近未接受治疗的学校,他们基于特定地理范围内总人口密度的差异在计量上识别了学校间的外部性。尽管正外部性导致理性人投资或消费不足在理论上很容易找到依据,但在实证上衡量外部性的程度和空间范围始终是个难题,驱虫实验提供了一个研究该问题的范本,这一实证方法被广泛用于后来的外部性研究中。

(三)环境依赖

随机对照实验的结果是否可以复制或推广至其他环境是一个引发争论的问题,这通常包含三个方面的顾虑(Duflo, Glennerster & Kremer, 2006):首先,项目实施是否过于复杂以至于难以复制,由于一些试点项目是在高质量的项目工作人员精心监管下实行的,非政府组织等实施机构是否能有效运行是值得怀疑的。其次,不同人群对相似项目的反应是否一致,即是否可以由某一人群对项目的反应预测另一人群对相似项目的反应相同,这并不仅限于随机评估,所有的实证研究均面临研究样本是否可一般化的问题。但随机实验通常实施于相对较小的地方而非使用国家的代表性数据集,这常常会加剧样本不具代表性的问题。最后,实验是否可以预测相似但不完全一样的项目的结果,即基于某一项目的特定影响,是否可以从中了解相似但不同的项目效果。

为减小实验的环境依赖,一方面可以通过重复实验提供不同地点、不同时间、不同样本的实验结果,例如驱虫项目后来被学者们不同环境下进行了重复实验。克雷默的驱虫实验实施于肯尼亚的农村小学, Bobonis, Miguel & Sharma(2006)在印度城市的幼儿园实施了补铁项目(当地贫血症高发),与研究者合作的非政府组织也不同,但对入学率、儿童身高、体重的影响表现出了惊人的相似。另一方面,研究者不可能用实验检验到每种情况下的结果,这就需要依赖于行为理论来判断某一情境下的项目是否会在另一情境下有效。因此,研究者可通过结合重复实验和理论来扩展外部有效

性。在不同版本的项目实施过程中,随着经验的积累,研究者可以推断出行为反应机制,结合基础理论框架便可提取出对于重复和推广项目最为重要的因素,获得一般化的结论,而即便该项目不能推广,实验所发现的个体行为模式也是具有一般性的。

另外,为保证项目扩展时的质量,在实验实施中清楚地记录项目中遵循的程序,并收集有关项目执行情况的数据,也可以帮助后续推广者清楚地了解评估的内容。

(四) 霍索恩效应和约翰·亨利效应

推广评估结果的另一个障碍是实验本身会使实验对象改变行为,使得研究者观测到的实验组和对照组的差别并不完全代表项目影响(Duflo, Glennerster & Kremer, 2006)。个体知道自身处于评估中而会做出的相应行为反应,其中实验组的行为改变被称为霍索恩效应(Hawthorne effect),对照组的行为改变被称为约翰·亨利效应(John Henry effect)。实验组可能因为感激或意识到被观测而改变在实验期间的行为,例如额外提供学校投入可能暂时鼓舞了学生和老师的士气;对照组可能因为未得到帮助并因此改变行为,例如对照组个体可能产生与实验组个体竞争的意识,或者因为沮丧而懈怠。

在实验项目外收集长期数据可以帮助评估个体行为的改变是观测效应还是项目本身的影响,迪弗洛等(Duflo & Hanna, 2006)实施了用照相机监测教师出勤情况的项目,在实验结束后,与研究者合作的非政府组织决定继续推行这一项目,一年以后,研究者发现该项目的影响并没有因为有了评估监测而改变,而是与初始评估阶段时的影响相似,这表明没有霍索恩效应。另外,也可以通过实验设计的方法加以甄别,例如在实验组和对照组之外,增加仅接受部分实验处理的组别(Ashraf, Karlan & Yin, 2006)。在评估某种特殊的存款项目是否可以刺激个体进行储蓄时,考虑到储蓄的增加有可能源于项目人员的介绍而非该存款产品的出现,研究者在实验组外设置了两个对照组,其中一个对照组不接受任何处理,另一个对照组仅接受宣传促销但无法购买该储蓄产品,从而可以将产品与宣传的影响分离开。

五、结论与启示

克雷默的实验研究方法为解决贫困问题提供了新的思路和证据,直接或间接地影响了发展中国家的减贫政策和项目。其设计的提高入学率、教育改革、改善学生健康等项目给了政策设计者以启发,并提供了小范围内实验效果的精准评估。他的研究成果可与各地区具体背景结合,丰富改革者可选择的有效政策集合,促使其优化机制设计,合理配置资源。

克雷默的实地实验不仅可以帮助政策制定者找出有效的扶贫政策,还可以帮其避免将资源浪费在无用的项目上。例如,此前世界银行曾将对肯尼亚的部分贷款用在了课本提供项目,但克雷默等(Glewwe, Kremer & Moulin, 2009)的实验指出,以提高学生成绩为目标增加教材提供或者其他投入效果均不显著,因此世界银行可以将资金转向对扶贫更有效的项目;再比如,肯尼亚的学校曾采取分班的方式试图改善学生学习效果,但克雷默等(Duflo, Dupas & Kremer, 2015)的实验发现,单纯减小班级规模并不是有效的办法。实地实验可以指导政策制定者更合理地运用有限的资源,节省了大量的时间和精力。

严格而系统的实验评估具有很强的外部有效性,了解某项项目有效或无效的收益不仅局限于实施该项目的国家或组织,还可以为其他正在寻找可行社会政策或援助计划的政府、非政府合作组织等机构提供指导。基于实验结果,已证明成功的项目可以在国内扩大规模甚至被其他国家和地区采用,而失败的项目可以被放弃。确定项目是否有效也可以帮助国际机构消除关于援助效果的质疑,促使其为落后地区的发展提供长期支持。

基于随机对照实验发现的经验也会以项目之外的其他方式间接影响真实政策的制定。例如,克雷默等(Kremer & Miguel, 2007)的研究并不仅仅意味着其他发展中国家是否应该照搬驱虫治疗项目,实验发现的贫困人口对预防性健康产品的需求价格弹性也具有很强的政策意义。基于价格会极大影响低收入人群进行健康投资的发现,英国政府曾引用这一研究呼吁贫困国家取消预防性健康产

品和服务的收费,很多发展中国家也确实基于该证据逐步取消了这一费用。另外,克雷默等发展的实验方法也在很大程度上影响了政府组织和非政府组织的工作,以克雷默等推广的科学研究方法为基础,兴起了很多致力于减轻全球贫困、评估政策提议的组织,他们对新项目采用随机对照实验的方法加以评估,并将所发现的证据切实用于现实项目中。

总之,克雷默将实验与发展问题相结合的研究极大地推动了减贫实践的发展。自从在肯尼亚实施的一系列关于教育的实验开始,克雷默将宏伟的发展计划拆解为具体可实施的政策目标,然后进行现实干预,评估不同措施的成本有效程度,丰富了减贫措施的可选择范围和人类对全球发展问题的认知,提高了我们减轻贫困问题的能力,也激励了大量研究者进入这一领域的研究。

在着眼于现实问题的同时,克雷默开创性的实验方法为发展经济学提供了一个新的强有力的工具,后续研究者可以此为基础研究更为广泛的问题,发展经济学进入了为发展政策提供经验支持的繁荣阶段。这一方法也成为经验研究的一个可依赖的武器,弥补了非实验实证方法的不足,并将实验经济学发展至空前的高度。此外,克雷默将经济理论与现实紧密结合起来,以理论为指导探究解决发展难题的思路,并以实践中的证据反过来启发经济学者重新思考理论。他所发展的实地实验重构了实证研究与理论检验之间的联系。此前的实证研究常是非实验的回溯性研究,即基于收集既成事实下的过往数据,评估恰巧已实行的项目,理论被用来支撑所检验假设的合理性。克雷默则反过来由理论出发寻找现实中可提供检验的证据,通过实地实验提供了一种预期性评估方式,将待检验的假设置于重要位置,通过灵活的实验设计加以验证。因此,克雷默可以由理论框架决定施行何种实验,随机对照实验作为一种研究工具大大提高了实证工作者检验和质疑理论的能力。

参考文献:

- Angrist J., E. Bettinger, E. Bloom, E. King & M. Kremer(2002), "Vouchers for private schooling in Colombia: Evidence from a randomized natural experiment", *American Economic Review* 92(5):1535—1558.
- Bertrand, M., E. Duflo & S. Mullainathan(2004), "How much should we trust differences-in-differences estimates?", *Quarterly Journal of Economics* 119(1):249—275.
- Bobonis, G. J., E. Miguel & C. Puri-Sharma(2006), "Anemia and school participation", *Journal of Human Resources* 41(4):692—721.
- Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan & F. H. Rogers(2006), "Missing in action: Teacher and health worker absence in developing countries", *Journal of Economic Perspectives* 20(1):91—116.
- Committee for the Prize in Economic Sciences in Memory of Alfred Nobel(2019), "Understanding development and poverty alleviation", Document of Nobel Prize in Economic Sciences.
- Crépon, B., E. Duflo, M. Gurgand, R. Rathelot & P. Zamora(2013), "Do labor market policies have displacement effects? Evidence from a clustered randomized experiment", *Quarterly Journal of Economics* 128(2):531—580.
- Duflo, E. (2006), "Field experiments in development economics", In: R. Blundell et al(eds), *Advances in Economics and Econometrics: Theory and Applications*, Cambridge University Press.
- Duflo, E., P. Dupas & M. Kremer(2011), "Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya", *American Economic Review* 101(5):1739—1774.
- Duflo, E., P. Dupas & M. Kremer (2015), "School governance, teacher incentives, and pupil-teacher ratios: Experimental evidence from Kenyan primary schools", *Journal of Public Economics* 123:92—110.
- Duflo, E., R. Glennerster & M. Kremer(2006), "Using randomization in development economics research: A toolkit", MIT Department of Economics Working Paper No. 06—36.
- Duflo, E. & R. Hanna(2006), "Monitoring works: Getting teachers to come to school," NBER Working Paper, No. 11880.
- Duflo, E. & M. Kremer(2003), "Use of randomization in the evaluation of development effectiveness", Paper prepared for the World Bank Operations Evaluation Department (OED) Conference on Evaluation and Development Effectiveness.
- Duflo, E., M. Kremer & J. Robinson(2008), "How high are rates of return to fertilizer? Evidence from field experiments in Kenya", *American Economic Review* 98(2):482—488.
- Duflo, E., M. Kremer & J. Robinson(2011), "Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya", *American Economic Review* 101(6):2350—2390.

- Fisher, R. A. (1925), *Statistical Methods for Research Workers*, Edinburgh: Oliver and Boyd.
- Glewwe, P., N. Ilias & M. Kremer(2010), “Teacher incentives”, *American Economic Journal: Applied Economics* 2(3):205–227.
- Glewwe, P., M. Kremer & S. Moulin(1998), “Textbooks and test scores: Evidence from a prospective evaluation in Kenya”, Natural Field Experiments 00255, The Field Experiments Website.
- Glewwe, P., M. Kremer & S. Moulin(2009), “Many children left behind? Textbooks and test scores in Kenya”, *American Economic Journal: Applied Economics* 1(1):112–135.
- Glewwe P., M. Kremer, S. Moulin & E. Zitzewitz(2004), “Retrospective vs. prospective analyses of school inputs: The case of flip charts in Kenya”, *Journal of Development Economics* 74(1):251–268.
- Holmström, B. (1999), “Managerial incentive problems: A dynamic perspective”, *Review of Economic Studies* 66(1):169–182.
- Kremer, M. (2003), “Randomized evaluations of educational programs in developing countries: Some lessons”, *American Economic Review* 93(2):102–106.
- Kremer, M., C. Brannen & R. Glennerster(2013), “The challenge of education and learning in the developing world”, *Science* 340(6130):297–300.
- Kremer, M., J. Lee, J. Robinson & O. Rostapshova(2013), “Behavioral biases and firm behavior: Evidence from Kenyan retail shops”, *American Economic Review* 103(3):362–368.
- Kremer, M., J. Leino, E. Miguel & A. P. Zwane(2011), “Spring cleaning: Rural water impacts, valuation and property rights institutions”, *Quarterly Journal of Economics* 126(1):145–205.
- Kremer, M. & E. Miguel(2007), “The illusion of sustainability”, *Quarterly Journal of Economics* 122(3): 1007–1065.
- Kremer, M., S. Moulin & R. Namunyu(2003), “Decentralization: A cautionary tale”, Natural Field Experiments 00290, The Field Experiments Website.
- Kremer, M., G. Rao & F. Schilbach(2019), “Behavioral development economics”, In: B. D. Bernheim et al(eds), *Handbook of Behavioral Economics: Foundations and Applications*, Vol. 2, Elsevier.
- Miguel, E. & M. Kremer(2004), “Worms: Identifying impacts on education and health in the presence of treatment externalities”, *Econometrica* 72(1):159–217.
- Schultz, T. P. (2004), “School subsidies for the poor: Evaluating the Mexican Progresa poverty program”, *Journal of Development Economics* 74(1):199–250.
- Vermeersch, C. & M. Kremer(2005), “School meals, educational achievement, and school competition: Evidence from a randomized evaluation”, World Bank Policy Research Working Paper, No. 3523.

Michael Kremer’s Contributions to Development Economics

ZHANG Yan WANG Qi
(Peking University, Beijing, China)

Abstract: Michael Kremer is awarded the Nobel Memorial Prize in Economics in 2019 with Abhijit Banerjee and Esther Duflo for their experimental approach to alleviating global poverty. Kremer introduces randomized controlled trials into research on poverty alleviation in Kenya. He provides empirical evidence for a large number of development policies guided by economic theory, and proposes many effective measures to alleviate global poverty. His research methods reshape the research paradigm of development economics, which make important contributions to the development of experimental economics and even the entire field of empirical research. This article focuses on Kremer’s experimental research on education, health, behavioral bias and other aspects to sort out and summarize his academic achievements, in order to understand his pioneering contributions in practical problems and research methodology.

Keywords: Michael Kremer; Randomized Controlled Trial; Development Economics; Poverty Alleviation; Nobel Prize

(责任编辑:李仁贵)

(校对:刘洪愧)