

走向真实世界的实验经济学

——田野实验研究综述

罗 俊 汪丁丁 叶 航 陈叶烽*

摘 要 本文对于经济学田野实验的研究方法、核心思想、发展脉络、分类界定进行了细致梳理,进而系统全面地介绍了近年来田野实验在四个重要的经济学科领域的应用实例,并展望了田野实验在中国的应用前景。以期为国内经济学界更好地认识田野实验研究,在复杂性和多样性并存的中国情境下,利用田野实验研究方法开展面向真实世界的经济学研究提供可能的研究思路和有益的建议。

关键词 实验室实验, 田野实验, 中国情境

一、引 言

近十年来,在经济学的实证研究领域,田野实验(field experiments)(或称随机性(控制)实验、随机性田野实验、randomized controlled experiments)一词正被越来越多地提及。由于应用领域上的局限,20 世纪 90 年代中期以前田野实验研究在经济学权威期刊上的发表数量并不多。¹而后随着田野实验研究人员对于实验方法的推进及应用领域的拓展,运用田野实验方法的研究在权威期刊上的发表数量逐年增长(如图 1),这一研究工具已被主流经济学界所接受,并已成为经济学领域非常重要的研究方法。这里有三项代表性的证据:一是据 Card *et al.* (2011)的统计,在过去的六年里,平均每年有 8—10 篇田野实验研究实例的论文发表在经济学的 Top5 期刊上;二是介绍田野实验的标志性论文——“Field Experiments”(Harrison and List, 2004)至今在 Google 学术搜索上的被引次数已超过千次²;三是 Esther Duflo 就因为将田野实验方法引入发展经济学而被授予 2010 年的克拉克奖。

* 罗俊、叶航、陈叶烽,浙江大学经济学院、浙江大学跨学科社会科学研究中心;汪丁丁,北京大学国家发展研究院、浙江大学跨学科社会科学研究中心。通信作者及地址:罗俊,浙江省杭州市浙大路 38 号浙江大学经济学院,310027,电话:18768161869;E-mail:roger.1228@163.com。本文受国家社科基金重点项目“关于新兴经济学理论创新的综合研究”(13AZD061)资助。作者非常感谢两位匿名审稿人的宝贵修改意见。

¹ 据 Card *et al.* (2011)的统计,1975—1994 年间,共有 12 篇田野实验的应用研究发表在经济学 Top5 期刊上。

² 截至 2013 年 7 月 1 日,该文被引用次数为 1175 次。

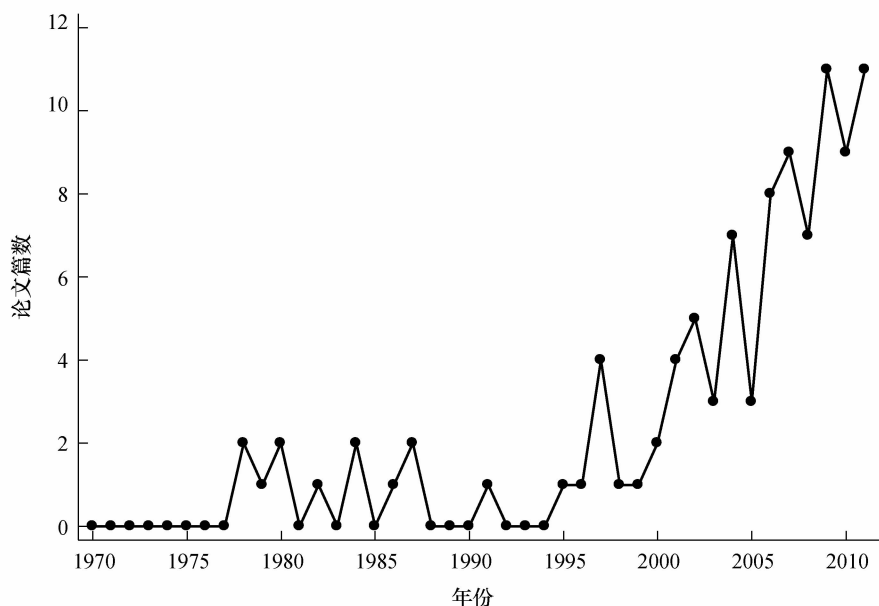


图 1 1975—2011 年在经济学 Top5 期刊上发表的田野实验研究文献数量
注：该图由 Card *et al.* (2011) 图 1 修改补充而来。

而与之形成鲜明对比的是，在国内经济学期刊上利用田野实验的研究方法所做的实证检验可谓凤毛麟角³，可以说，国内经济学界对于田野实验方法的使用还处于谨慎观望的态度。当然，这也与经济学的田野实验研究正是方兴未艾之时，与之相关的概念、名称较为丰富和繁杂⁴，且尚未有明确、公认的界定有关。鉴于此，我们认为充分详细地对田野实验研究方法的发展脉络、核心思想、分类界定进行梳理，系统全面地介绍近年来田野实验在几个重要的经济学科领域的应用实例，可以促进国内经济学界更好地认识田野实验研究；帮助非实验经济学领域的经济学者掌握田野实验这一实证研究方法，并将其应用于各分支学科领域；也可以拓展实验经济学者设计田野实验实例的研究思路，并在复杂性和多样性并存的中国情境下，以贴近现实著称的田野实验研究方法为开展面向真实世界的经济学研究提供更多的机会和可能。

因此，本文将分六个部分对经济学田野实验做全面的综述。第一部分为引言；第二部分从两个方面对田野实验的发展历程进行系统的梳理；第三部分对田野实验的概念、特点及与其他实证研究方法的比较做具体的总结；第四部分对田野实验在行为经济学、劳动经济学、发展经济学、公共

³ 据我们所知，何浩然(2011)、何浩然和陈叶烽(2012)是仅有的两篇发表在国内权威经济学期刊上的关于田野实验的实证研究。

⁴ 如与田野实验含义相近的名词(随机控制实验、项目评估方法、实验室实验、社会实验等)就有十余种之多。

经济学四个领域的应用实例做细致的综述；第五部分介绍田野实验在中国的应用前景，并陈述在中国情境下如何更好地开展田野实验研究的思路；第六部分为结语。

二、田野实验的发展历程

田野实验同其他研究方法一样，从有争议到被接受，从分立到合流，经历了一个不断完善吸收的过程。总体来看，可梳理出两条独立的田野实验发展轨迹，即在经济学实验室实验基础上关于现实情境的补充而催生出的田野实验；在实验方法上对社会实验的改进而产生的田野实验。

（一）贴近现实：从实验室到田野

田野实验的一个重要方法论来源是实验经济学家对实验室实验(laboratory experiments)研究方法的反思。由 Smith(1962)等人发展的实验室实验研究方法于 20 世纪中叶逐渐兴起，并在 21 世纪走向繁盛，现已成为一个重要的经济学研究领域和数据来源。

经济学实验室实验吸收了自然科学实验中的“控制”，但人的行为总是会被某些不可控的因素影响。以研究社会偏好(social preference)的经济学实验为例，实验中人的行为除了受到物质激励的影响外，还非常容易受到其他多种因素的影响：(1)道德伦理和社会规范的约束；(2)实验员的注视；(3)决策时嵌入的情境；(4)样本选择的代表性问题；(5)博弈的初始禀赋(Levitt and List, 2007)。这些因素的影响使得实验经济学家对实验室实验结论外推至真实世界产生了一定的疑问，并开始共同推动实验经济学另一个方向的发展，即田野实验。田野实验与实验室实验相比，突出的特点就是更贴近真实世界，正如 Carpenter *et al.* (2005)发出的感慨：“实验经济学家不再保守，他们开始在实地(field)招募被试而不是实验室；他们开始用实际物品而不是诱导价值(induced valuation)；他们开始在实验说明中设定真实的情境而不是抽象的术语。”

需要注意的是，这里对于实验室实验的诘问并不意味着否定实验室实验这一研究方法。⁵实际上，实验室实验仍旧是帮助经济学家更科学地检验经济理论、分析人的经济行为的重要工具。实验室实验与田野实验在很多时候可以互补。实验经济学家常在同一批被试中开展田野实验和实验室实验，以检验情境因素是否可以影响人的行为。可以认为，田野实验是对实验室实验的

⁵ 我们也不认为田野实验可以取代其他实证研究方法，而只是提供了另一种有效的选择。正如 Falk and Heckman(2009)所说：“田野数据、调查数据、实验室实验、田野实验以及标准的计量方法可以共同推动社会科学的研究。”

一般性结论在具体情境下的补充和检验。Harrison and List(2004)因此根据实验中相关情境因素引入的程度而将田野实验划分为不同的类型。

表 1 列举了各种田野实验之间及与实验室实验（conventional lab experiments）、自然实验⁶（natural experiments）、非实验数据相比较的差别。其中实验室实验招募在校大学生作为被试，设定抽象的背景和实验规则；人为的田野实验（artefactual field experiments）除不限于招募在校大学生作为被试以外，其他均与实验室实验一致；框架的田野实验（framed field experiments）中交易物品、实验任务、被试代人的实验信息均是真实的情境，其他与人为的田野实验一致；自然的田野实验（natural field experiments）中实验在自然发生的环境中展开且被试不知道自己正在参与实验，其他与框架的田野实验一致。

表 1 实验分类及相关比较

	被试 类型	被试 信息	激励	实验 环境	被试知 晓实验	研究者 干预	外生 改变
实验室 实验	大学生	抽象 信息	诱导 价值	实验室	是	是	否
人为的 田野实验	各类 人群	抽象 信息	诱导 价值	实验室	是	是	否
框架的 田野实验	各类 人群	情境 信息	真实 激励	自然 环境	是	是	否
自然的 田野实验	各类 人群	情境 信息	真实 激励	自然 环境	否	是	否
自然 实验	各类 人群	情境 信息	真实 激励	自然 环境	否	否	是
非实验 数据	各类 人群	情境 信息	真实 激励	自然 环境	否	否	否

注：该表参考了 List and Rasul(2011)表 1 的内容。

（二）方法的改进：从社会实验到田野实验

如果说实验室实验的发展催生了田野实验的兴起，那么社会实验可认为是田野实验的前身。准确地说，田野实验是经济学家对社会实验在实验方法上的加强和改进。

Fisher(1925)和 Neyman(1933)运用实验方法解决农业生产领域的工作常被认定为经济学田野实验最早的雏形。虽然这个阶段的田野实验没有一项研究对象是人，且很少发表在经济学期刊上。但他们的研究确实为以后的田野实验的设计和样本选取方面提供了基本的思路。其中，Fisher 在实验设计中

⁶ 自然实验在很多特征上与自然的田野实验相似，如同样是对自然发生的事件进行考察，也具有一定的实验属性（如随机分配被试），但这类实验是由外生的政策干预所造成的，缺乏人为的实验控制（Harrison and List，2004），因此在检验干预的效果时往往会存在潜在的内生性问题。

引入了随机性的概念并且强调了实验的三要素：复制、干扰和随机性。Neyman 意识到重复随机抽样的作用，并认为概率推断的必要条件是随机性。这些研究被认为是第一次将随机性这一实验方法的关键要素概念化了(Levitt and List, 2009)。

20 世纪后半叶是田野实验研究发展的第二个重要时期，彼时政府机构开展了一系列大规模的社会实验。在欧洲，早期的社会实验包括 60 年代末英国电价的设计；在美国，社会实验能够追溯到关于个人所得税项目检验的工作。其他一些大规模社会实验项目包括住房补助等。这些实验对于当时政策的确定有着重要的影响，比如它们对于 1988 年家庭支持法案⁷的贡献，就修正了救助无自理能力儿童(AFDC)这一计划。

这一阶段的社会实验虽然为当时的社会公共项目的实施提供了一定的实践依据，但其在实验方法上却存在诸多缺陷和隐患。这些缺陷主要包括以下几个方面：一是随机偏误(randomization bias)。随机分配被试进入实验可能改变参与者的行为。例如被分配到实验组的被试可能会更加努力，进而影响项目的评估效果。二是中途退出(attrition)实验。由于社会实验历时较长，一些被试可能在实验过程中就离开了当地。三是样本的非代表性(non-representative sample)问题。比如项目参加者可能为自愿报名的义工，而义工的平均素质一般高于普通人，因此造成实验被试的自选择效应。四是小型实验的条件与项目被大规模推广时的现实条件不同。由于条件的制约，一些小型实验在大规模推广时可能会遇到一些困难和差异，而这种差异又可能导致对项目的评估结果截然不同。

面对社会实验在实验方法上的不足，近来开展的优秀田野实验都有针对性地进行了一些加强和改进。譬如，虽然都是通过随机招募被试来参加实验，但与社会实验不同的是，这些田野实验都尽可能在自然发生的环境中随机招募普通的人群，以期使得作为研究对象的被试都没有意识到自己成了实验的一部分。此外，这一时期的田野实验也总是比社会实验有着更大的理论目标，这些田野实验在许多情况下是被设计去检验经济理论的，包括收集构建理论的有效信息、组织数据以测度理论假设下的关键参数等。

田野实验的两条发展轨迹的侧重点与方法论来源都有所不同，所应用的实际领域也有所不同。从实验室到田野的演变路径，关注行为的社会情境属性，因此这类田野实验多应用于行为经济学与劳动经济学；从社会实验到田野实验的演进路径，多是对宏观领域问题的检验，因此这类田野实验多应用于发展经济学和公共经济学。图 2 对田野实验的两条发展轨迹及应用领域做

⁷ 1988 年美国实施的家庭支持法案旨在针对不负责任的父母强制执行子女抚养命令，以及通过扩大职业训练和教育的机会来降低父母对于抚养费及社会福利的依赖程度。该法案在一定程度上改革了美国的福利体系。

了概括。⁸

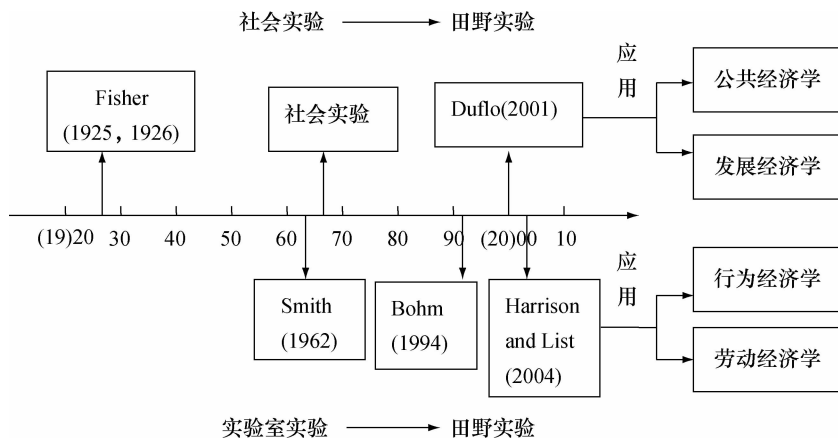


图2 田野实验发展轨迹及应用领域

三、田野实验的概念、特点及相关实证方法比较

(一) 田野实验的核心思想、概念和基本操作方法

通过对田野实验发展历程做的梳理，可以勾勒出田野实验的核心思想、概念和基本操作方法。从研究内容上来说，田野实验的核心思想在于对现实世界所发生的事件之间的联系给出科学的检验；从研究方法上来说，田野实验的核心思想在于利用实验的操作技巧(如随机化被试、控制相关变量)，来评估干预事件的处理效应(treatment effect)或者说变量之间的因果效应。

在此，可以给田野实验做一个概念上的总结：运用科学的实验方法去检验真实世界中(或者说自然发生的环境下)，而不是在实验室里发生的扰动对人们行为决策的因果影响，它可以为经济学理论研究提供有力的实证依据。

田野实验的基本操作方法是从一个总体中随机选取被试样本，然后随机将被试分为控制组与处理组，在控制其他因素不变的情况下，对处理组被试进行实验处理，并根据随后两组被试的数据比较，得出最后的因果效应。由于被试是被随机分入两组的，因此实验的处理组水平完全独立于个体特征和其他可能影响实验结果的因素。这就避免了计量模型中常见的遗漏变量偏差(omitted variable bias)或内生变量偏差(endogeneity bias)的问题，即处理组

⁸ 其中 Peter Bohm 是最早对实验室实验和田野实验的差别做系统性阐述的经济学家，并于 1994 年的文章中首次在实验经济学界引入 field experiments 这一名词。但可惜的是他已于 2005 年去世，且其对田野实验的应用也仅限于环境、农业等领域，因此他在经济学田野实验发展历程中的地位一直以来都没有被广泛认识。直到 Dufwenberg and Harrison(2008)的文章中对于 Peter Bohm 相关实验研究的专题介绍，并追认 Peter Bohm 为田野实验之父，人们才开始认识到 Peter Bohm 对田野实验发展的重要贡献。

水平 X 对实验结果 Y 的因果效应表现在条件期望上的差别, $E(Y | X=x) - E(Y | X=0)$ 。其中 $E(Y | X=x)$ 是处理组被试在处理组水平为 x 时的行为表现 Y 的期望值, $E(Y | X=0)$ 是控制组被试的行为表现 Y 的期望值。

(二) 田野实验的特点

正如我们在以上对田野实验的核心思想和概念的叙述中所反复提到的, 田野实验这一实证研究方法的主要特点, 其实就是现实性和科学性。

1. 田野实验的现实性

现实性指的是田野实验的实验过程贴近真实世界。具体而言, 相对于实验室实验的大学生被试来说, 田野实验的被试更具针对性和广泛性。他们可以是车间工人 (Carpenter *et al.*, 2005)、公司 CEO (Fehr and List, 2004)、渔民 (Carpenter and Seki, 2011) 等。已有研究也发现大学生被试与这些特殊群体被试在行为偏好上确实存在一定的差异, 如车间工人在实验中比大学生表现得更慷慨、更公平 (Carpenter *et al.*, 2005); 公司 CEO 在实验中也比大学生表现出了更高的信任度和可信度 (Fehr and List, 2004)。

除被试选择的针对性和广泛性以外, 实验设计所嵌入的真实情境也是田野实验现实性的重要体现。这里的真实情境不是封闭的实验室, 而大多是日常生活中的场景和环境, 如聚餐时的付账方式 (Gneezy *et al.*, 2004)、求职市场中筛选简历 (Bertrand and Mullainathan, 2004) 等。同一批被试在嵌入真实情境的田野实验与抽象的实验室实验中, 可能会表现出不同的行为偏好。Gneezy *et al.* (2004) 招募了一批被试分别参与一个田野实验和一个实验室实验, 田野实验的设计是考察被试在不同付账模式下的消费情况。此外, 他们要参与一个有着相似实验结构的实验室实验, 只是真实的消费情境变成了在不同的付费机制下为抽象的物品支付禀赋。结果表明, 实验室的行为数据 displays 更多社会偏好, 但田野实验中被试的行为却更多地与经典的经济学理论相吻合。

2. 田野实验的科学性

田野实验的另一个重要特点是科学性, 主要体现在相对其他实证分析方法而言, 能更为直接和便利地对变量之间的因果关系做出检验。检验变量之间的因果关系不仅是经济学实证分析的主要目的, 也是所有科学研究的基本目标之一。对因果关系的推断必须是基于反事实 (counterfactual) 框架的 (Heckman, 2008), 即一个影响变量或者干预对个体的因果效应, 应该是个体在控制组和干预组这两个可能的结果状态之间的差异。但就像“人不能两次踏入同一条河流”一样, 个体的结果只可能在一个组中被观察到。这样, 对于该个体而言, 干预结果与控制结果中必有其一是缺失数据。⁹ 因此, 对因

⁹ 谢宇,《社会学方法与定量研究》,北京:社会科学文献出版社,2006年。

果关系的科学分析只能用平均干预效应来替代,但替代的前提是,干预组和控制组必须其他相关因素上是一致的。

而田野实验解决这一问题的办法,是由随机分配将实验被试分成实验处理组(即干预组)和控制组,以实现“其他相关因素都无显著差异”的条件,从而通过比较两组被试的表现,就能最终确定干预对被试表现的因果效应。但是,在非实验条件下进行因果关系的有效识别却并不易做到,研究者往往需要搜集大量的自然数据,利用复杂的统计方法和计量模型才能实现对变量之间因果关系的检验,如以下几种常见的与因果检验相关的非实验研究方法。

(1) 格兰杰因果关系检验(Granger causality test): 格兰杰因果关系检验的基本观念在于,未来的事件不会对目前与过去产生因果影响,而过去的事件才可能对现在及未来产生影响(Granger, 1969)。在时间序列情形下,两个经济变量之间的格兰杰因果关系可定义为:在控制了 y 变量的过去值以后, x 变量的过去值仍能对 y 变量有显著的解释能力,就能认为 x 变量是引致 y 变量的格兰杰原因(Granger-cause)。但是,格兰杰因果关系和真正的因果关系还是有本质区别的。前者说的是一种时间先后顺序的可预测性,而真正的因果关系是一种逻辑上的顺序。

(2) 工具变量(instrumental variable, IV): 工具变量的基本思想是,遗漏变量(omitted variable) u 的存在使得直接利用 x 与 y 的数据做计量回归,会出现内生性(endogeneity)的问题,即 x 对 y 的影响可能是由于 u 的存在所导致的,而采用 z 这一工具变量来替代性地检验 x 对 y 的因果影响则可以避免以上问题,图 3 是工具变量功能示意简图。但工具变量法的使用至少需要满足两个条件: z 与 u 不相关,即 $\text{Cov}(z, u) = 0$; z 与 x 相关,即 $\text{Cov}(z, x) \neq 0$ 。¹⁰ 因此要找到这样一个符合条件的工具变量,并不是一件容易的事情。

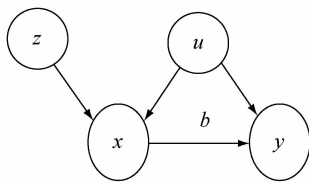


图 3 工具变量功能示意图

(3) 倾向值评分分配比法(propensity score matching, PSM): 在实证研究中,存在很多其他变量混淆自变量和因变量之间的因果关系,混淆变量的影响通常被称为选择性偏误(selection bias),而通过倾向值匹配的方式可在一定程度上控制和消除选择性偏误。首先,倾向值是指在控制可观测到的混淆变量的情况下,被研究的个体受到自变量影响的条件概率(Rosenbaum and Ru-

¹⁰ 伍德里奇,《计量经济学导论:现代观点》,费剑平译。北京:中国人民大学出版社,2010年。

bin, 1983)。配比的基本逻辑是将受到自变量影响的个体与没有受到影响的个体进行配对, 而倾向值匹配就是保证匹配起来的个体的倾向值相等或者近似。以检验大学教育对个人收入的作用为例, 倾向值匹配就是将没有受过大学教育的人和受过大学教育的人配对并保证他们的倾向值(即上大学的概率)相同或近似, 这样就得到了“上过大学”和“没上过大学”两组个体。因为混淆变量已经在基于倾向值的配对过程中被控制起来了, 两组个体收入上的差异就只能归因于大学教育的有无, 而不是其他混淆变量, 由此遏制了选择性误差。¹¹可见这一方法的理念就是通过对非实验数据进行一定的处理以实现类似于实验数据的随机控制效果。

(4) 双重差分模型(difference in difference, DID): 双重差分估计方法利用了个体受到事件干预前后的数据, 即控制时变性因素影响的同时, 通过比较受到事件干预影响的处理组和没有受到干预影响的控制组的数据差异, 来最终检验事件对个体干预的因果效应, 图4为双重差分模型示意图。设定 Y_{c1} 为控制组被试在干预前的平均数据, Y_{c2} 为控制组被试在干预后的平均数据, Y_{t1} 为处理组被试在干预前的平均数据, Y_{t2} 为处理组被试在干预后的平均数据。因此, 由图4可知干预的因果效应在数值上等于 $(Y_{t2} - Y_{t1}) - (Y_{c2} - Y_{c1})$ 。但这一数值上的相等至少需要满足两个条件: 对照组个体的数据确实未受到政策影响; 对照组个体受时变性因素影响的幅度与处理组必须一致, 即二者在坐标轴中的斜率必须相等(Harrison and List, 2004)。

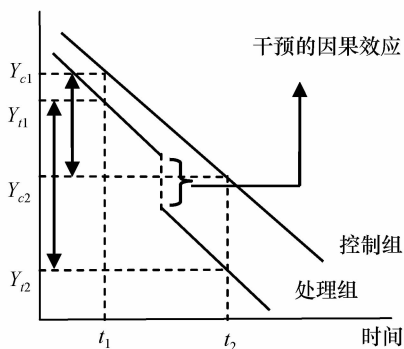


图4 双重差分模型示意图

根据以上所介绍的几种与因果检验相关的非实验研究方法的特点, 我们很容易发现这些研究方法在检验变量之间的因果关系时, 要么不能检验逻辑上的因果关系, 要么必须满足多个苛刻的条件, 要么操作过程复杂且最终仍是为了实现类似于实验数据的随机控制效果。因此, 可以说田野实验在科学地检验自然情境下变量之间的因果关系时, 有着明显的优势。

¹¹ 胡安宁, “倾向值匹配与因果推论: 方法论述评”, 《社会学研究》, 2012年第1期, 第221—246页。

四、田野实验的具体应用

在这部分内容中,我们将全面地介绍近十年来发表在权威期刊上有代表性的田野实验研究实例,这些研究实例都是田野实验方法在行为经济学、劳动经济学、发展经济学、公共经济学这四个重要经济学科领域的应用。

(一) 行为经济学

1. 利他行为

近十几年来,迅速发展的行为经济学领域的一个重要发现是违背经济人自利假设的亲社会行为的广泛存在¹²,这些行为主要包括利他、公平、信任、互惠、合作等。然而实验经济学家在观察到这些亲社会行为存在的同时,也常常在田野实验中发现这些行为的表现并不稳定,如人们的利他行为就很容易受到捐赠信息的公开、社会压力等社会性因素的影响。

Soetevent(2005)就研究了教堂里捐献的匿名与否对利他捐献行为的影响。为了分析捐献的匿名性对于被试捐献行为的影响,实验设计了两种不同的收集捐款的容器,实验的控制组用一个密封的小包募捐,处理组则是一个开放的篮子。区别在于,使用篮子作为捐款容器时,周围的捐款者均能观察到自己的捐款数量,而且每个捐款者在捐款时均能看到篮子里已经捐款的数量。实验结果表明,当篮子作为捐款容器时,这一非匿名性的效应可以显著增加捐款的数量。

DellaVigna *et al.* (2012)利用田野实验证实了社会压力对个人捐赠行为的作用。实验的主体任务是上门募捐,包含三个形式各异的实验组:基准组为直接上门募捐;实验组1的工作人员会于募捐前一天留下传单,以告知被试第二天将会有人上门募捐;实验组2的传单上有个选项是“若您不想被打扰,可以在方框内打钩”,如果第二天工作人员在传单上发现被试选择不希望被打扰,则不会敲门募捐。研究者报告了以下结果:发送传单不会对平均捐赠率有太大影响,但当传单上印有可以选择不被打扰的选项时,捐赠率则会比基准组减少30%,说明一般家庭都会尽可能避免募捐,这一结果与社会压力假设一致。

2. 公平行为

田野实验研究常常关注到被试所代入实验的背景信息会如何影响他们互动交往时的行为表现,这一信息主要包括被试在社会情境下所遵循的行为规范和制度。Carpenter *et al.* (2005)通过比较学生与工人在分配博弈中的行为

¹² 国内关于亲社会行为和社会偏好检验的系统研究可参见陈叶烽等(2012)。

表现来测度社会情境对公平行为的影响。实验采用的是最后通牒博弈和独裁者博弈，被试是在堪萨斯城的车间工人、明德学院的大学生和堪萨斯城社区学院的大学生。三组实验被试均在各自所处的现实环境中参与实验。实验设计通过比较车间工人和明德学院的大学生的实验行为数据体现了两个维度的效应：社会情境与被试的社会人口特征。实验结果发现，堪萨斯的学生在最后通牒博弈中比堪萨斯的工人提供了更多的分配，然而在独裁者博弈中，车间工人则比两所学校的学生提供了更多的分配，且车间工人在两种博弈中的分配额没有显著的改变，这意味着工厂车间的社会情境使得工人的行为趋向于更公平和更多分享的表现。

3. 信任行为

如前所述，实验室实验被试为在校大学生，其中重要的不足是忽略了不同年龄段人群的行为差异。Sutter and Kocher(2007)的田野实验对不同年龄段的群体信任度(trust)和可信度(trustworthiness)进行了测度。实验者招募了不同年龄组别被试参与信任博弈(trust game)实验。结果发现，从儿童时期到青少年时期，人们对他人的信任度几乎呈线性增长，而进入成年阶段后，不同年龄组别的信任度却没有显著不同。此外，人们可信度则存在于所有年龄组别中，并随着年龄的增长而显著增加。

除去年龄这一个体属性会影响人们的亲社会行为表现以外，职业属性对亲社会行为表现的影响也在田野实验研究中被检验。Fehr and List(2004)比较了哥斯达黎加的CEO和大学生在信任博弈中的表现。实验设计了两组处理任务，一组是标准的信任博弈，另一组则在信任博弈中加入了惩罚机制，委托人可以选择对没有达到回报要求¹³的代理人给予固定筹码的惩罚。最终的统计结果显示，CEO比大学生表现出了更高的信任度和可信度。而博弈中委托人可以对代理人施加惩罚这一举动也带来了回报：如果委托人在可以施加惩罚的情况下对代理人表现出了信任，那么代理人将比没有惩罚这一环节的博弈中表现出更高的可信度。

4. 合作行为

实验室实验中的被试在做行为博弈时通常是被抽离了各自社会身份的，而田野实验则可以关注到社会身份的引入对被试行为的影响。Ruffle and Sosis(2004)考察了人们在面对组群内部和组群外部对象时合作行为的差别。实验招募的被试来自以色列特有的集体合作农场基布兹(Kibbutz)以及周边的一些现代城镇，基布兹人仍然保持着集体农业生产组织形式，内部实行“各尽所能，平均分配”的分配原则。实验设计类似于公共品博弈(public good game)实验，两组实验任务分别为：在两个匿名的基布兹人中开展，在一个公

¹³ 委托人要求代理人的准确回报数是不会影响结果的，因为委托人的惩罚不需要成本，所以委托人总是会选择惩罚的。

开身份信息的基布兹人和一个公开身份信息的城市人中开展。实验结果显示基布兹人在与另一位基布兹人的配对中比与城市人的配对中表现出了更高的合作水平,而且在基布兹人与城市人的配对中,两者表现出的合作水平没有显著差异。

田野实验还可以应用在研究虚拟网络社区的合作行为上。Chen *et al.* (2010)设计了一个田野实验,来探究社会比较的作用能否增加网络社区的总体参与水平。研究者给网络用户发送了两种不同的信息:用户评价数量的中位数、群组中用户平均的净分值,控制组用户接收到的信息是用户自己过去的评分行为。随后研究者会跟踪用户在收到信息后一个月内的网络操作,结果发现在收到用户评价数量中位数信息的用户中,在中位数以下的用户在一个月内会有 530% 的电影评价数量的增加,在中位数以上的用户会有 62% 的评价数量的减少;而在收到用户平均净分值信息的用户中,平均值以上的用户在这个月内会从事一些帮助其他用户的行为。

5. 其他个体行为偏好

除以上关于社会偏好的研究以外,田野实验在行为经济学方向上的应用还涉及对人们在交易中的非理性偏好、个体的时间偏好、风险偏好的测度等。如 List (2003)通过田野实验发现有市场经验的个体的交易行为会更接近于新古典经济学假设中的理性人假设。研究者在真实的运动卡片市场中招募被试进行市场交易。被试包括有卡片市场交易经验及没有卡片市场交易经验的两类人员,他们被随机分配在不同禀赋商品的实验组。市场交易程度由交易率、成交价与商品的价值等实验数据表征。结果表明,市场中确实存在禀赋效应 (endowment effect),但禀赋效应会随着被试市场经验的增加而逐渐消失,即个体的市场经验可以消除市场的异常 (market anomalies)。

Harrison *et al.* (2002)在丹麦开展的大范围田野实验中,通过带有真实物质激励的调查问卷测度了不同群体在不同投资期限内 (6 个月、12 个月、24 个月和 36 个月) 的折现率水平。在实验任务 1 中,被试随机参加一场关于四种投资期限中的一种折现率的问卷调查,在实验任务 2 中,被试需要对四种投资期限的折现率调查都做出选择,但实验员只以其中一种折现率调查的选择作为被试的实验报酬来支付。实验结果指出,被试 12 个月投资期到 36 个月投资期的名义折现率都是保持不变的,但不同社会经济变量的个体之间却有着显著不同的折现率。

田野实验研究常会将个体时间偏好的测度和人们在现实生活中的储蓄意愿结合起来讨论。Ashraf *et al.* (2006)利用菲律宾一家银行设计承诺性储蓄产品的契机开展了实验,实验被试是银行的既有客户,他们被随机分配在了两个任务组,推荐给实验组中的被试是一款承诺储蓄产品,被试一旦开户就只有在账户达到规定的期限或金额后才能撤回资金,但无须承诺一定要在账户上存入金额。控制组中的被试则会被鼓动向既有的账户存入金额或没有接

收到任何干预。而在此之前，所有被试还要参与一个关于时间偏好的问卷调查，研究者以此来判断个体的时间偏好类型。6个月后，研究者们发现实验组的被试平均储蓄额比控制组高出47%，此外，在有关时间偏好的调查问卷中，女性表现出了更低的时间折现率，即更偏好于长期的承诺。

（二）劳动经济学

1. 市场歧视

劳动力市场中关于性别、种族等歧视多有发生。基于统计数据回归的检验为歧视的普遍存在提供了经验证据，但由于无法控制其他没有观察到的影响生产率的变量，因此难以进一步解析歧视的性质及原因，而田野实验的方法为这一设想提供了可能。Bertrand and Mullainathan(2004)开展的自然的田野实验考察了美国劳动力市场中的种族歧视问题。实验员首先通过人口普查数据和调查问卷的方式，找到较为常见的且被人们普遍认为是某一肤色(黑人、白人)人种中特有的姓名，进而在虚构的求职者简历中随机安排对半的这两类求职者姓名，并控制除种族以外的其他可能影响求职的变量。此外，实验员还构造了两类质量不同的简历以检验简历质量会如何影响不同种族求职者在接收面试通知数量上的差异。结果表明，投出的白人简历收到的面试通知比黑人简历要多50%，白人简历质量的提升可以增加2.5%的面试通知，而黑人简历却只有0.51%的增加。

List(2004)在真实的运动卡片买卖市场中开展的田野实验研究对歧视的性质做了深入分析。这里的歧视是指市场的双边交易中常出现的对于弱势群体(女性、非白人、老年人)的价格歧视。实验员通过分别招募买(卖)方被试、控制买(卖)方人群属性，并将其带入实验发生的卖(买)方市场中，以检验卖(买)方对于购买(售卖)相同商品的不同属性人群是否会提供不同的价格。实验结果表明针对不同人群的价格歧视确实存在，市场中真实的交易者面对弱势群体时会提供更劣等的价格。¹⁴同时，买方所遭遇的价格歧视相比较卖方会更明显。

2. 员工激励

考察工人在不同激励机制下的反应是激励理论的重要内容，但由于内生性的问题，利用企业层面的数据计量不同激励机制对于工人表现的影响时可能遗漏了其他没有观察到的因素，从而不能准确地识别出激励效应。而通过实验的方法，在外生给定不同的激励机制下测度生产率，则可以直接解决内生性的问题。Shearer(2004)在林业公司开展的田野实验就比较了工人在计件工资制和定额工资制下的生产率。实验员随机选择了一些工人参与实验，每

¹⁴ 即面对弱势群体的卖方，买方会提供更低的价格；面对弱势群体的买方，卖方会提供更高的价格。

个工人开始被随机安排在计件工资或定额工资的薪酬体系下,在被观察 60 个工作日后,又被轮换安排到另一薪酬体系下观察 60 个工作日。实验结果表明,工人在计件工资的薪酬体系下的平均生产率比定额工资的薪酬体系下高出 20%。

还可考虑针对管理者的激励会如何影响生产率的问题。Bandiera *et al.* (2007)在企业开展的田野实验,外生地改变管理者的激励机制,在固定工资的基础上增加绩效工资,进而通过观察工人们每个工作日的表现来测度对管理者的付酬方式会如何影响企业的效率及工人生产率的差距。实验的主要发现如下:(1)管理者绩效工资的引入增加了工人的平均生产率(21%)和工人之间的生产率方差系数(38%)。(2)生产率的个体数据揭示出能力最强的工人的生产率有显著的提高,然而其他工人的生产率则没有受到影响甚至有所下降。此外,管理者在引入绩效工资后,工作能力最强的工人更多地被管理者招入自己的工作组中,而工作能力最弱的工人则更少地被管理者招入。(3)生产率更高的工人会受到更多的关注,进而更可能留在工作组中。被招入工作组的工人生产率更高,进而会受到更多的关注。因此,实验的发现表明管理者激励与工人收入不平等之间存在交互关系,从实证上检验了激励机制对企业或个人绩效的因果关系。

3. 雇佣关系

在新古典经济学中的均衡状态下,企业应该付给工人市场出清时的工资,工人则提供最小化的努力。而在礼物交换模型中,企业付给工人超出市场出清时的工资,工人也提供更多的努力。后来,这一假设也在实验室实验中得到了证实(Fehr *et al.*, 1993)。然而,有疑问的是,招募被试模拟雇主与工人,被试在电脑前的决策作为雇主给定的工资水平和工人提供的努力程度,这样的抽象模拟能否真实反映劳动力市场中的决策行为。鉴于此,Gneezy and List(2006)在现实的劳动力市场中进行自然的田野实验来检验礼物交换模型。实验招募大学生为大学图书馆录入书籍信息。招募时告知这一工作为一次性工作,工作将持续 6 个小时,每小时报酬为 12 美元,被试并不知晓自己正在参与一个实验。实验分为两个任务组,任务组 1 的被试按照事先允诺的报酬完成整个过程,任务组 2 的被试在开始工作前被告知报酬将由之前的每小时 12 美元提高到每小时 20 美元。实验结果显示,在刚开始的一段时间里,雇员在更高的薪酬下比更低的薪酬下付出了更多的努力,然而在这段时间以后,这一努力的不同将不再被观察到,即礼物交换模型只在开始的一段时间内得到了验证。

在标准的经济学框架中,雇员都是理性的欺骗者,当他们认识到偷懒的边际回报超过了边际成本,便会机会主义地选择卸责。而企业应对的办法是依靠监督和激励机制来改变卸责的收益。利用非实验方法对上述理论进行检验时会遇到两个困难:首先,卸责行为本身难以测量,且卸责行为可能会与

监督力度存在双向因果关系；其次，在考虑卸责行为与监督力度的关系时可能遗漏其他没有观察到的因素，如雇佣制度的影响等，这些影响都可能造成计量时的内生性问题。

Nagin *et al.* (2002)设计的田野实验控制了其他可能的影响变量，通过外生改变对雇员机会主义行为的监督力度来克服以上检验理性欺骗者模型的困难。实验在一家电话邀约公司进行，电话接线员的薪酬按计件工资支付，由其成功完成的电话邀约数量决定。但是，管理者对最后捐款是否落实存在不完全信息，雇员有激励谎报自己的邀约数量。为制约这一机会主义行为，管理者的监督是对一部分成功邀约到的电话进行随机回访。实验结果发现，有一部分的雇员表现如“理性欺骗者模型”中所展示的，那些平日里认为管理者对其不够公平和关注的雇员，在有力的监督下会显著地减少谎报行为。另一方面，也有部分雇员不论管理者监督力度如何，都不会显著改变其行为。这意味着管理者在执行针对雇员机会主义行为的监督时，应充分考虑雇员个体的异质性。

4. 家庭决策

主流经济学关于行为决策的理论模型中，常常忽视影响家庭内部决策的因素，大多数模型要么将家庭视作一个个体决策者，要么认为家庭决策是成员在完全信息和充分交流下通过讨价还价过程形成的一致决定。但对于家庭决策更为现实的假设是夫妻之间可能存在有限的交流和私有的信息。在这一假设下，考虑夫妻之间的信息交流对双方决策的调整及最终家庭决策的影响，可以完善既有理论模型。当然，这一因果关系需要外生变量的引入才能更好地辨别。Ashraf(2009)利用田野实验辨别了信息和交流是如何影响家庭金融决策的。实验在菲律宾的农村银行中进行，被试为银行客户。他们需要决定如何支配被赠予的初始禀赋，是消费还是存在自己或家庭的账户上。实验有三个不同条件的任务组，任务组1(私有信息组)中夫妻双方分别进行决策，互相之间不能交流，不知晓对方是否获得禀赋以及做了怎样的决策。任务组2(没有事前交流的公开信息组)中夫妻双方知晓对方的收益，且同时做决策。决策过程中不能相互交流，也无法观察到对方的决策。任务组3(公开信息组)中夫妻双方知晓对方的收益，且在正式决策之前可以互相交流，也能观察到对方的决策。

有三个主要的实验结论：(1)私有信息组中，男性更多地选择将钱存入自己的账户，而在没有交流的公开信息组中，男性则更多地选择消费；(2)那些家庭存款由丈夫控制的家庭中女性的决策行为，与那些家庭存款由妻子控制的家庭中男性的决策行为非常相近；(3)决策前的交流使得大多数男性被试将钱存入了妻子的账户，而不是选择消费或存入自己的账户。实验结果显示信息与交流的改变能在很大程度上影响家庭决策。

(三) 发展经济学

发展经济学家如今也越来越多地寻求与政府、非营利机构、企业等的合作,以评估随机设置下的社会干预(social intervention)实验所带来的效应。干预实验通常被设计去解答具体情境下的实际问题,如怎样让学生更积极地来上学,如何帮助当地居民更多地储蓄,怎样使父母愿意带孩子来接种疫苗,有什么有效的办法能够遏制当地的腐败问题等。

1. 储蓄问题

Dupas and Robinson(2013)在肯尼亚开展的田野实验,解释了为什么穷人的储蓄水平被抑制。研究者随机安排被试在五个有着不同设置的实验组(一个控制组和四个任务组)。所有实验组的个体都被鼓励为了个人健康去储蓄,并在实验之前为自己设置一个健康目标,但只有任务组的个体会被提供相应的储蓄设计以帮助他们达成目标。实验任务组 1 提供了一个明确、安全的地方存放资金。实验任务组 2、实验任务组 4 与实验任务组 1 的不同之处在于,这两个任务组中被试的储蓄只能分别用于预防性健康产品或突发性健康产品,而实验任务组 1 中被试的储蓄可以同时用于这两种产品。而实验任务组 3 则表征了社会承诺和信用对特定用途储蓄的作用。这里的社会承诺和信用来自民间集资信贷联盟中其他组成员的社会压力。

实验结果显示,所有储蓄产品的采纳率都很高,最低的有 66%,最高的有 97%,这是由于所有的实验设计都为储蓄提供了安全的场所和明确的目的,这一提高人们储蓄水平的机制设计似乎是一种心理账户(mental accounting)¹⁵,被试在实验后的调查问卷中普遍报告说,一旦资金被预留为了特定目的,他们一般就不会再把钱花在其他未列入日常计划的事宜中。对于大部分人来说,对预防性健康产品的投资是无效的,因为人们认为预防性的健康产品所绑定的流动性成本太大,而针对突发性健康产品的储蓄设计则提高了人们在这一方面的投资。另外,通过信用和社会压力来促进组内成员为健康产品所做的储蓄,是一种非常有效的提高人们在健康产品上的投资方式。

2. 腐败问题

众所周知,腐败是导致发展中国家经济发展缓慢的主要原因。但即便这一问题如此重要,由于直接测度腐败活动的困难,理论界至今对如何更好地遏制腐败问题仍然鲜有一致意见。Becker and Stigler(1974)认为结合监督与惩罚可很好地控制腐败的发生,但监督者自身也可能出现腐败,所以自上而下的监督能否有效地解决腐败问题仍然没有确定的答案。另一个可能的解决途径是社会民众的参与,即增加底层对于当地项目的监督。但对当地项目这一

¹⁵ 心理账户这一概念主要指的是人们心理上对结果的编码、分类和估计的过程,即人们会把现实中的客观等价的支出或收益在心理上划分到不同的账户中。

公共品的监督,也可能要面对民众搭便车的问题。因此,要判断这一方案是否可行也仍然需要通过实证的检验。

为比较这两类解决途径,Olken(2007)在印度尼西亚的农村开展了相关的田野实验。实验利用的是印尼全国性农村大建设之机,每个农村都需要修建一条乡间公路。作者在建设款项到位之后、项目正式动工之前,随机地选择了一些乡村告知其将接受中央审计机构的审计,最终的审计结果将在公开的乡村大会上宣布,若有腐败问题将接受社会的制裁。为了检验民众参与的监督所起的作用,作者设计了两个不同的实验。第一个实验中村民被邀请参与审计大会的监督,旁听官员汇报项目资金使用情况。第二个实验中审计大会开始前,村民需要匿名填写一张关于项目执行和资金使用情况的评论表格,评论汇总情况将在大会上宣布。为评估这些监督干预措施对于腐败问题的作用,还需对腐败本身进行度量。在项目完成后,作者通过检查公路样本对项目原材料的质量进行评估,通过调查和面访原材料供应商及工人以对采购价格和人工费用进行核算,最终估计出项目实际所花费用,并与官员所声称的项目花费进行对比,相差的部分用来衡量官员在项目中的腐败问题。根据实验数据发现,引入自上而下的外部监督可以显著地减少项目资金的流失(8%),而有民众参与的监督则不能显著地减少项目资金的流失。

3. 教育问题

教育所带来的回报在多大程度上影响了学校的决策,学生本人是否清楚教育可能给他们带来的回报?一些民众可能由于低估了教育的回报而导致了低就学率。因此,只是为人们提供教育高回报的信息,也许就能极大地改善发展中国家的教育问题。

Jensen(2010)在多米尼加共和国开展的社会干预田野实验中检验了以上假设。作者做了长达四年的跟踪调查,以期考察这一干预的长期效应。作者首先通过面板数据发现,大多数的学生会在八年级(义务教育的最后一年)时选择辍学,且他们普遍认为接受高中教育带来的回报很低(而实际上回报是很高的)。为了验证这一观念是否导致当地的高辍学率,作者随机选择给一部分在校学生提供当地劳动力工资数据信息,另一部分不提供信息的学生为控制组。实验结果发现,在六个月后的再次访谈中,被提供信息的学生对于教育回报的判断有显著的提高。在四年后的调查中,被提供信息的学生所接受的平均教育年限比控制组的学生多出0.2年。作者在对个体异质性数据做进一步分析后发现,尽管信息的提供使得几乎所有学生对于教育回报的判断有显著提高,但那些最穷的学生却仍然没有因此提高就学率,那些相对更富的学生则平均增加了0.33年的教育年限。

发展经济学家们在考虑如何提高学生就学率的同时,也意识到就学率的提高并不一定能给学生带来最大的收益,而相关的援助项目也可能没有起到提高学生成绩的作用。Banerjee *et al.* (2007)认为如果把项目的援助对象锁定

为学习上更需帮助的学生,也许能使项目的投入发挥更大的作用。为此,作者开展了两个旨在帮助学生提高学习成绩的社会干预实验。第一个实验是专门针对落后学生的教育改良项目,由来自当地社区的年轻妈妈在课外辅导三、四年级学生语文和数学的学习,每天两小时左右。第二个实验是面向所有四年级学生的计算机技能学习项目,学生们通过完成计算机中的数字游戏来提高数学的计算能力,每周两小时左右。

数据显示,这两项社会干预实验都显著地提高了学生的学习成绩。其中,所有接受教育改良项目的学校学生在第一年里,与控制组学校学生的平均成绩的方差增加了 0.14,在第二年里方差增加了 0.28;所有接受计算机学习项目的学校学生在第一年里,与控制组学校学生的数学成绩的方差增加了 0.35,在第二年里方差增加了 0.47。

4. 医疗问题

Miguel and Kremer(2004)检验了在肯尼亚开展的为儿童除虫计划对当地教育和儿童健康的作用。以往类似的研究是通过对个体层面的随机化处理,来比较接受了治疗的处理组被试与未接受治疗的控制组被试在之后的考试表现上的差异,以此估计儿童除虫计划对发展中国家教育的贡献。但这类研究的问题在于忽视了接受治疗的处理组被试对于减少疾病传播的外部有益性(externality benefits)而可能低估了除虫计划的整体作用。此篇文章的不同在于,作者从学校层面进行了实验的随机化处理。实验结果表明除虫计划显著地提高了处理组学校的到课率。在外部有益性方面,除虫计划也显著地提高了处理组学校与邻近学校中未接受治疗的儿童的健康及参与教育的积极性。

为发展中国家提供医疗补助的项目,总是基于必要的医疗物品能够为当地社会带来正的外部性,从而提高总体社会福利的目的。但项目还要考虑成本问题,对补助项目的医疗物品收取一定的费用,是否可以有效地节省在并不是真正需要或不使用该类物品的当地居民上的支出,进而提高公共项目的效益。有田野实验表明,当公共项目的物品从免费到收取一定费用时,居民对这一物品的需求将急剧下降到正常需求以下(Kremer and Miguel, 2007)。此外,那些因为支付不起物品价格的贫困居民却可能是最需要这类医疗物品的人。为了具体估计以上物品的需求弹性,辨别医疗物品免费援助及分摊成本的举措为当地居民带来的外部性,Cohen and Dupas(2010)在肯尼亚开展供应防虫蚊帐¹⁶项目的田野实验。作者随机选择了 20 家诊所来供应蚊帐,其中 4 家为控制组,其余 16 家以 0(免费分发)到 40 先令(肯尼亚货币,40 先令相当于 0.6 美元)的四组价格随机供应。¹⁷为探知那些最需要蚊帐的孕妇是否愿意支付费用,作者对所有孕妇的血红素进行检测,依此判断孕妇的健康状况,

¹⁶ 这一物品可以被用来防止产妇感染,进而有效地减少产妇贫血和新生儿死亡率。

¹⁷ 防虫蚊帐的市场价格为 400 先令。

并在蚊帐发放数月后对领取者进行家访，以检查蚊帐是否正在被使用。

实验结果显示，当物品价格从0增加到0.15美元时，需求并没有显著的下降，而当物品价格从0增加到0.6美元时，需求则下降了60%。没有证据显示对蚊帐的支付意愿与使用需求相关，虽然那些出价最高的孕妇比免费领取的孕妇更多地使用了蚊帐，但是那些在中间价买到蚊帐的孕妇却更少地使用了蚊帐。没有证据显示物品从免费领取到成本分摊会使得那些更需要蚊帐的妇女购入蚊帐，因为通过血红素检测发现，那些愿意支付更高价格的妇女并不比控制组被试身体更弱。因此，作者根据需要使用弹性的参数估计，结合个人和社会的回报(婴儿死亡率)认为免费援助与成本分摊方式一样是有经济效益的。

5. 其他问题

除了以上我们提到的腐败问题、教育问题、医疗问题以外，田野实验的研究方法在发展经济学领域的应用还涉及发展中国家的民众监督与公共卫生服务的实施、扶贫项目与信贷市场、小额信贷项目、微型企业扶持计划等问题。

如何有效监督和落实在发展中国家实施的公共项目，一直是理论界和政府部门关注的重要问题。Björkman and Svensson(2009)在乌干达开展的田野实验考察了通过鼓励民众参与监督的方式，能否有效地保证公共卫生服务的实施，从而改善当地居民的健康状况。此外，经济学家们还关心发展中国家的公共项目落实后，能为当地带来多大程度的收益。Angelucci and De Giorgi(2009)联合墨西哥农村扶贫项目设计的田野实验，发现对相关家庭的资助可以间接地带动当地的消费，并激活保险和信贷市场。因此也认为为避免低估该类项目在改善当地状况上的效用，应从村落层面进行随机化设计，而不仅是对接受项目的被试与同一村落的其他被试进行比较。

小额信贷项目在发展中国家的应用是发展经济学新生的研究领域，田野实验在这一领域的应用也有很多研究成果。Karlan(2005)利用实验对秘鲁小额信贷的偿款做了研究，作者在借款人中开展了信任博弈和公共品博弈实验，并获得了这些借款人在乡村小额信贷中的个人借贷、储蓄、偿还率等数据。实验目的是检验行为博弈实验中所考察的被试的合作、信任行为能否用来预测其真实的借贷行为。Karlan and Zinman(2008)在南非开展的田野实验，检验了小额信贷项目的借贷者对利率不敏感的假设，发现借贷者的贷款数额更多地是与贷款期限相关，而不是与利率相关。De Mel *et al.* (2008)开展的田野实验中，随机给斯里兰卡的微型企业发放补助以调查企业真实的资本回报率，进而分析了企业的异质性(企业家能力、家族财富、风险偏好等)对于资本回报率的影响，为当地小额信贷的实施提供了有效依据。

(四) 公共经济学

田野实验方法在公共经济学领域的应用,有的是对随机选定的被试设计不同的作用机制,来比较不同的作用机制下被试的行为数据,以考察不同机制在解决社会公共问题时的效果;有的是通过比较在相关公共政策影响范围内的人群和控制组人群的表现差异,来检验公共政策的效应。

1. 慈善募捐

List and Lucking-Reiley(2002)较早地研究了慈善募捐中的激励机制,如本金额(seed money)和返还制度(refund policy)的作用。实验通过向当地居民募集资金来为环境政策研究中心购置设备。实验设计了不同水平的本金额(即分别已经准备了总募集资金3000美元的10%、33%、67%,余下的比例需要向居民募集)和返还制度(若募集资金没有达到目标额将返还所捐资金)以比较不同情况下的捐助数量。实验结果揭示,随着本金额从10%到67%的增加,居民集资的参与率和平均捐助额都有显著的增加,而返还制度的设立也使得平均捐助额有显著增加。

Landry *et al.* (2006)分别检验了募捐的初始本金额、单奖金的彩票捐款、多奖金的彩票捐款对个人捐款的作用。实验是以为防治灾害研究中心募捐的名义开展,设计了四个不同的任务组进行比较,分别是有初始本金的自愿捐款(募捐者会被告知中心已收到1000美元的匿名募捐者的捐款)、无初始本金的自愿捐款、单奖金的彩票捐款(每捐助1美元将获得一张彩票,并有机会获得1000美元(1名)的大奖)、多奖金的彩票捐款(每捐助1美元将获得一张彩票,并有机会获得250美元(4名)的大奖)。结果显示,有彩票的捐款机制比自愿捐款机制的收益高出50%,参与率也高出100%。

2. 退休储蓄计划

Duflo and Saez(2003)借助田野实验方法对美国递延税收¹⁸(Tax Deferred Account, TDA)退休计划做了讨论。人们对于项目计划相关信息的掌握,可以决定自己在退休账户中做多少储蓄。提供了金融知识培训的企业,其员工也会更多地参与到退休储蓄计划中来。然而,企业提供的这一知识信息可能是内生的(个人选择参与项目可能来自企业的社会网络效应),因而使得结论的因果关系颇为复杂。

为克服这一困难,Duflo and Saez(2003)通过实验分析了信息提供与社会交往在学校职员加入TDA计划中所起的作用。作者随机选择给一些还没有加入TDA计划中的被试发放邀请函,请他们参加TDA计划说明会,为他们提供有关计划的信息,参会者将获得20美元的出场费。实验结果表明,在这一

¹⁸ 为提高企业及其员工储蓄养老的积极性,美国国内税法规定,企业员工向退休储蓄账户的缴款在规定限额内可在税前扣除,享受延迟纳税待遇。

过程中, 接收到邀请函的被试参加说明会的人数是控制组被试的5倍, 没有收到邀请函而来参加说明会的人数也是控制组被试的3倍, 这显示邀请函不仅提高了个人参与的积极性, 还在他们的同事中产生了外溢的社会效应。此外, 作者还在说明会后调查了被试参与计划的情况, 结果显示说明会的召开可以有效地增加人们参与TDA计划的可能性, 但收到邀请函与否并不显著影响最终参与TDA计划的可能性。实验说明一些小的物质激励、社会交往和对计划信息的知晓容易影响人们对退休储蓄所做的决定。

还有实验研究对美国另一项重要的退休储蓄计划——401(k)¹⁹的实施效率做了考察和优化。Carroll *et al.* (2009)利用企业改变401(k)储蓄计划中登记制度的机会, 通过自然田野实验的方法, 分别在两个时间段对两批员工实施了两种登记制度。为比较“积极决策”(企业要求员工需在受雇后的30天内反馈自己是否愿意参与401(k)计划)与“传统登记制度”(登记时默认的选择是不参与401(k)计划)这两种登记制度对401(k)计划的参与率和效率的影响, 研究者控制了两批员工的储蓄偏好, 并确保其他可能对员工参与401(k)计划造成影响的因素在两个时间段是基本一致的。实验结果表明, “积极决策”制度下, 员工参与401(k)计划的比率比“传统登记制度”下多了28%。此外, “积极决策”的登记制度还提高了401(k)计划的平均储蓄率和累积结余。

3. 青少年成长项目

有许多分析认为, 居住地周边同龄人的素质高低极大地影响青少年的行为。但由于在邻里属性的测度及同龄人群界定上的不确定, 使得对同群效应(peer effect)的估计存在一定的困难。而理想的解决办法是随机安排家庭在不同类型的居住地, 以比较在不同环境下长大的青少年的行为。Katz *et al.* (2001)研究了美国MTO(Moving to Opportunity)计划²⁰, 在随机实验的方式下测度邻里属性对弱势家庭生活机会的影响, 并评估住房迁移计划的有效性。实验将家庭随机分配到以下三组: (1)实验组: 接受住房优惠券, 迁往贫困率不到10%的地方居住, 并为该组家庭提供迁移咨询, 以帮助这些家庭在规定时间内租赁到合适的房屋; (2)Section 8计划组²¹: 接受住房优惠券但不限制去处, 有时间限制, 但不接受任何迁移咨询; (3)控制组: 不分配住房优惠券, 但仍然能接受公共住房的援助。研究结果显示, MTO计划对弱势家庭的孩子有显著的积极影响, 实验组和对比组的孩子均呈现出较少的行为问题, 且实验组孩子受伤害、犯罪的可能性都较低。

我们知道, 教育对于个人今后的成长和发展起着重要的作用, 但是否幼

¹⁹ 401(k)是美国的企业养老金计划。该计划规定, 企业需为每位员工设立专门的401(k)账户, 员工每月从其工资中拿出一定比例的资金存入养老金账户, 而企业也按一定的比例往这一账户中存入相应的资金。员工退休时即可从这一账户中领取资金。

²⁰ 美国政府为分散弱势群体, 实现社会各阶层混合居住的公共房屋迁移政策。

²¹ 指美国的住房与社区开发法案的第八部分政策。

儿教育也能在一定程度上起这样的作用呢? 美国的教育成就计划(The Student/Teacher Achievement Ratio, STAR)对幼儿园的儿童和老师开展了大规模的跟踪实验, 教师和学生随机被分配到了不同的教室(15 人的小班教室和 22 人的大班教室), 以考察幼儿教育对于个人今后成长和发展的作用。Chetty *et al.* (2011)在对实验数据进行分析后, 得出在小班接受幼儿教育的学生比在大班接受幼儿教育的学生更多地进入了大学, 在大学里也有更好的表现; 有经验的教师所带班级的学生在工作后获得了更高的收入; 在更好的班级环境下(由其他同学期末成绩衡量)成长的学生在大学入学率和工作收入上都显著更高。

4. 其他公共项目计划

Yoeli *et al.* (2013)与美国加州一家公共电力公司合作开展的田野实验是一个电力用户“需求反馈计划”活动。电力需求反馈计划的目的是更加合理有效地分配电力。该计划是用户自愿参加的, 这类似于一个公共品问题, 计划的顺利实施能增进总的社会福利, 但参加该计划会给用户带来不便。为了减少这种搭便车现象, 作者通过设计决策的可观察性来验证声誉效应的引入能否提高用户的参与度。

首先, 研究者将邀请住户参与计划的邀请函投递到社区公共信箱, 以征求住户是否参与计划的反馈。该实验设计分为可观察组和匿名组, 可观察组中被试得知参与住户的姓名和单元号将被公开张贴在社区公告栏; 匿名组中的公布名单上只有代码, 没有个人身份信息。实验结果显示可观察组被试参与该计划的概率是匿名组的近 3 倍。以上实验设计证明了可观察性大大增加了人们对于公共品的参与度。那么这种可观察性是否会在个人关系和声誉拥有更大影响力的群体里起更大的作用? 在公寓楼中的住户相比于在独栋别墅里的住户, 他们更多地与邻居接触, 所以个人声誉会有更大的效应。租客和房东相比流动性更大, 也更少与他们的邻居们打交道, 那么租客会更少地在乎个人声誉。实验结果也显示, 可观察性大大增加了公寓楼住户和房主的参与度, 但是对独栋别墅住户和租客均无显著影响。

Kling *et al.* (2012)针对美国联邦医疗处方药物保险计划²²(Medicare Part D prescription drug insurance)开展了一个随机的田野实验。保险计划的覆盖依赖于消费者的选择, 由于个体的异质性偏好, 服务提供方需设计不同的选择方案来匹配相应人群的偏好, 且提供所有方案的相关信息也是提高各方利益

²² 美国联邦医疗处方药物保险计划是美国联邦医疗保险的一部分, 这一计划从 2006 年 1 月 1 日起开始实行, 不管收入、健康状况或目前的药费如何, 每个人都可以自愿地选择这一计划。参与者可以有两种支付方案: 一种是标准的, 另一种是大病的。标准的方案平均每月需要支付 37 美元。看病时超过 265 美元的支付后, 联邦医疗保险会支付 75% 的药费, 直到每年 2400 美元的封顶线为止。超过封顶线后需要自付, 但当每年支付药费超过 3850 美元时, 大病的处方保险计划可以为投保者报销其余 95% 的药费。此外, 这一计划针对不同人群还会有各种相应的优惠方案。

的重要手段。虽然消费者可轻易获得各方案的信息，但并不能保证他们会充分利用这些信息，解决这一问题的办法是免费提供不同方案之间的比较数据，且让消费者意识到这些信息确实能帮助他们获益。为此，研究者设计了两个任务组：干预组和控制组，干预组中参与者会被提供按照他们个人的药物信息所计算的一些数据，包括所有保险方案下需要支付的成本，以及如果从现有方案转移到最低廉方案下可节省的花费。控制组中的参与者只是被提供了包含如上信息的网址。实验结果发现干预组中被试的方案转移率达到了28%，而控制组的这一比例只有17%，且干预组被试预计平均每人每年将节省100美元的花费。

五、田野实验在中国的应用前景

随着中国与世界的逐渐接轨，中国经济学界已能紧跟国际学术前沿的步伐，在模型构建与数据分析上越来越精深、细致。而另一方面，经济学的田野实验研究正在兴起，而中国这片有着广袤“田野”及特有文化背景的土地完全可以提供丰富的社会“实验场”。倚赖这一天然的优势，也许中国的经济学者可以在逼近高深的数理工具的同时，也能为经济学的田野实验这一前沿研究方法贡献更多有独创性的且贴近现实世界的研究实例。

（一）中国情境下的田野实验

中国情境下的田野实验在我们看来有两条研究思路。第一条思路是基于不同文化背景下中国人行为模式与西方人行为模式的可能差异，采用中国的被试按照经典的实验设计程序来运行实验，以检验中国文化背景下经典实验的结果是否会发生改变。Henrich *et al.* (2001)在五大洲12个国家15个经济和文化环境迥异的小规模社会，开展了跨文化行为博弈实验。而中国各区域间的经济、文化、习俗也有着明显的差异，若能参照以上模式以经济发展水平、教育普及率、民族、习俗等变量作为异质性指标，在全国范围内遴选相应的样本产生点开展大规模的行为博弈实验，则不仅可以在理论上检验各变量对于人们行为表现的影响，在实践意义上也可依此建立国人处理利益冲突、利益分享时的行为特征数据库。现在仅有的尝试是Chen and Tang(2009)在拉萨、厦门、新加坡三地开展的行为博弈实验，分离出了文化和宗教对被试行为表现的影响；何浩然(2011)²³分别对贵州省的大学生和农村居民的个人及家庭跨期决策行为进行了模型识别和实验比较。Gong and Yang(2012)对云南婆娑族人和彝族人风险偏好的实验研究。

²³ 何浩然，“个人和家庭跨期决策与被试异质性——基于随机效用理论的实验经济学分析”，《管理世界》，2011年第12期，第12—31页。

第二条思路则更具创新意义,指的是围绕中国社会经济转型期涌现出的独有现象而展开的经济学田野实验。类似的实验已经在印度、印度尼西亚等发展中国家开展,而涉及当代中国转型期社会现象的经济学田野实验则有 Dulleck *et al.* (2012) 利用礼物交换博弈对中国的户籍身份在劳动力市场的歧视情况进行了检验。Cadsby *et al.* (2008) 将中国人的“关系”概念引入了信任博弈的田野实验,以讨论社会距离对于中国人的互惠、信任行为的影响。Cameron *et al.* (2013) 通过开展一系列的行为实验比较计划生育政策前后出生的被试的实验数据,分离出了没有其他兄弟姐妹一起成长对于独生子女的性格、行为所造成的内在影响。

(二) 与组织、机构的合作

为贴近现实,田野实验不像实验室一样只招募在校大学生作为被试²⁴,而是面向社会招募被试;田野实验也不拘泥于教室、机房开展实验,而是开始走向社会去运作实验。但相伴生的问题是高额的实验费用,职业人群大多需要比在校大学生更高的物质激励才能诱导出可靠的行为表现,而在复杂的社会环境下也要比在教室、机房中运作实验需要更多的人力和物力。

因而,独立开展田野实验的难度可想而知。鉴于此,诸多的田野实验研究者无不选择与组织和机构进行合作。此举既可以支持研究者开展田野实验,也可以帮助政府部门、慈善机构、企业提高组织运行效率,以图实现双赢。如与政府部门合作考察政策制度和项目计划可能带来的影响²⁵;与产业组织合作设计拍卖等市场运行机制;与学校、公共事业组织等合作实施公共品供给的田野实验;与慈善机构合作研究人们的慈善捐赠行为可能受到哪些因素的影响;与企业、工厂合作探索工人的工作绩效受到什么因素的激励,并以此优化组织的激励机制等。

但以我们自己正在开展的田野实验的研究经历来看,国内的学术研究与组织和机构之间还缺乏日常的合作机制,组织和机构也往往不以学术研究的数据与结论作为决策依据。因此,调动组织和机构合作的积极性是在国内顺利开展田野实验的关键。还需特别注意的是,与组织和机构的合作也可能会由于合作方的制约而存在研究偏误的风险,有时研究者的目标与合作方的意图并不一致,田野实验的结果可能还是政府和企业不能接受的数据。在此情况下,研究者计划的实验任务很容易受到合作方的抵制。因此,为了避免可能因为合作产生的分歧,实验者需在有条件的情况下争取开展更多的相关实

²⁴ 实验室实验的被试常被认为不具有代表性。Henrich *et al.* (2010) 就把这类群体称为“WEIRD”的人群:即来自西方(western)、受良好教育(educated)、工业化(industrialized)、富有(rich)、民主(democratic)的社会。

²⁵ 实际上,很多发达国家的公共项目及发展中国家的扶贫计划在实施过程中,都会邀约经济学家开展相应的田野实验,以便科学地评估项目的有效性。

验以求数据的准确性，并增强与合作方的沟通、恪守必要的学术道德。

（三）数据获取

相对实验室实验来说，田野实验的一个主要弱点在于实验可复制性上的困难。实验可复制意味着其他研究者可根据实验设计对实验结果进行稳健性检验，以达到证实或证伪的目的。其中，自然的田野实验相较其他类型的田野实验来说，又更为不可复制。因为，自然的田野实验往往需要利用现实世界发生的事件，现实世界的瞬息万变使得它的发生常带有偶然性。

但田野实验的这一固有弱点或许可以通过数据获取方式上的补充而有所改进。田野实验主要采用的招募被试方式有信件、电子邮件、上门招募、网络招募、通过专业的调查公司。其中通过网络获取被试数据的方式，成本低、原始数据易保存，这都将有利于其他研究者对实验进行重复的操作和检验。另外，这类实验还可以方便地根据实验目的，进行相关的改变条件操作，以克服田野实验较难控制的弱点，且网络用户不会意识到实验的进行，从而避免被试在实验过程中因为研究者的注视而做出刻意行为。

因此通过网络环境来设计实验，获取数据的方式已经被越来越多的研究者所采用（如 Chen *et al.*，2010 等）。中国有着庞大的网民群体和快速发展的网络环境，可预见丰富的网络资源将为经济学田野实验在中国的开展提供更多的数据获取机会，这样得天独厚的外部条件也可以帮助中国的学者在田野实验研究的新方向上收获更优异的成果。

（四）跨学科的交流

经济学田野实验有大量检验人的亲社会行为的研究，且这一类实验往往延伸到小规模社会群体和相对落后地区。实际上，在这类社会和地区开展田野研究(field study)于人类学、社会学领域颇为常见。人类学家和社会学家进入实地，经历长时间与当地人的交往，采用观察、访谈等研究方法，以获取第一手关于研究对象的资料。

他们的研究以定性为主，缺乏实证工具作为支撑。因此，兼具现实性与实证性的田野实验研究开始受到人类学家的广泛关注和参与(Tracer, 2003)。这里，尤其需要提及的研究是 Gurven and Winking(2008)在玻利维亚村庄开展的田野实验，两位人类学家对被试在行为实验中表现的社会偏好及在日常生活中的田野数据(如食物分享、社会参与等合作互惠行为)进行了比较。如此经济学实验和人类学田野调查相结合的研究方式对于跨学科研究来说，是个有益的尝试。

而对于中国情境下的经济学田野实验来说，中国地大物博的优势将非常有利于比较地区间亲社会行为差异及小规模群体社会习俗的实例研究。此类田野实验需要深入中国本土的乡村田野，少数民族聚居地等受到现代性影响

较小的地区开展。在这些地区开展实验之前,必须对当地的风土民情、乡俗村规有足够的了解,这样才能因地制宜地展开被试的招募、实验说明的设计、实验任务的实施等实验过程。但经济学家们毕竟对于亲自深入偏远落后地区开展一线调查的工作不甚熟悉,这就要求我们与长期进行田野调查的人类学、民俗学、社会学的学者进行深入的合作交流,向他们学习如何做真正贴近现实世界的研究。

六、田野实验:科学化与走向真实世界

1590年,伽利略在比萨斜塔上做的自由落体实验,撼动了亚里士多德以降强调主观推理和逻辑演绎,只以定性和先验的讨论来认识世界的方式,并从此开启了物理学的实验科学之路。与自然科学一样,早期的经济理论在很大程度上也是经济学家通过内省与观察,再借助演绎假说的方法构建产生的。从亚当·斯密的“看不见的手”的自由市场理论到马歇尔的局部均衡分析,无一不是如此。但自20世纪以后西方科学哲学的转向,波普尔在判定科学时的“证伪主义”标准被广泛接受,经济学能否成为一门可被证伪的科学,已然成为衡量经济理论是否与客观规律一致以及学科能否进一步发展的关键。如此的学科发展要求在一定程度上催生出了经济学计量方法的使用,即从大量的自然数据中获得相关的经济结构信息,并通过统计检验和推断以求能够证伪理论。

然而,计量经济学本身在操作方法和数据获取上所存在的一些局限性,使得我们很难通过对自然数据进行计量分析,就能直接证伪或检验经济学理论或模型是否成立、何种情况下成立。如此一来,实验这一自然科学中常见的证伪方法在经济学的实证领域中逐渐兴起也就不足为奇了。实验者能通过实验过程中的人为调节和控制来排除其他无关因素的干扰,以使得经济理论能够在不断地被检验或被证伪的过程中得到发展。经济学的实验之路为经济学日益科学化的进程提供了坚实的方法上的支撑。

实验在为经济学理论发展提供科学验证工具以外,还在努力走向真实世界,力图为经济学在实际问题中的应用提供直接检验和具体判断。经济学田野实验的有效应用可以避免人们对经济理论普适性应用的质疑,并从经济学理论研究的一般性结论回归到了对现实问题的针对性解答,如田野实验在发展经济学领域的应用,可以改变以往发展经济学只注重宏观理论分析的倾向,而将着眼点放在了一个个促进发展中国家或地区进步的具体项目上。经济学田野实验的有效应用还可避免人们对于“黑板经济学”的诘问,从以往经济理论对于人的行为的抽象分析转向到对带有现实情境的人的行为的研究,如田野实验在行为经济学领域的应用,可以克服自然数据回归对行为模型设定的依赖以及实验室实验中对于被试行为检验的外部有效性的不足。

田野实验更接近现实世界的属性意味着其实验结果是人们在一个特定情境下的经济行为表现，这也就决定了我们不可能将一个独立的田野实验结果视为一般性的结论，而由于实验者在田野实验中难以做到如在实验室实验中的自如控制其他干扰因素，因此一个独立的田野实验也往往不能直接做到对相关理论的检验和区分(Levitt and List, 2009)。

虽然田野实验本身并不是一种与理论联系非常紧密的实证研究工具，但其对于理论却仍具有重要意义：一是虽然单个田野实验不容易得到一般性的结论，但它至少可以通过一个自然情境下的实验结果来预测其他相近情境下人的行为表现，“我们可以预见小规模田野实验在其他相似环境中的运用，就如联邦体系中新的法律法规的补充往往是先来自于相关制度在地方州政府的成功试验”(List, 2007)；二是实验经济学家其实常常会通过田野实验之间的比较来得到有价值的一般性结论，如 Henrich *et al.* (2001) 从不同经济、文化社会的人们在田野实验中的行为表现差异，来得到经济或文化各自对行为的单一影响，这在某种程度上也是一种“保持其他条件不变”的控制手段；三是田野实验与其他实证工具(实验室实验、调查问卷数据等)的结合可以很好地验证理论、估计模型参数。如 DellaVigna *et al.* (2012) 关于捐赠的田野实验研究，是先通过在被试实验后所填写的调查问卷中获得的其在日常生活中的机会成本及弹性数据，才能最终根据被试在田野实验中的捐赠表现，估计出捐赠的社会压力模型中各参数。

田野实验除以上所述对于理论的意义以外，其重要作用还在于对现实问题的直接解答，比如市场中是否有歧视发生，歧视的性质是什么？为什么人们会有慈善举动，什么机制可以更好地促进慈善捐助？怎样的教育改革制度是最有效的？什么样的公共政策能降低当地青少年犯罪率？面对这些应用性课题，田野实验都展示了比其他实证方法更能直接有效地提供事物之间因果关系的特性，从而回答事物是否确实已然发生，以及它为什么会发生。

当然，任何研究方法都不可能白璧无瑕。田野实验的开展也有其局限和不足之处，这就要求我们在实际研究中，一方面，尽可能有效地实现田野实验的控制性和保证今后实验的可重复性²⁶；另一方面，我们也要明晰田野实验的角色，作为经济学实证研究中的一种工具，它并不可能独立解决所有理论与应用的衔接问题，这需要田野实验与自然数据、调查数据、实验室实验、计量分析等其他实证研究方法共同协作努力才有可能实现。

²⁶ 如最大限度地确保实验全部原始数据的可保存和可公开，以及实验全过程的可复制(在条件允许的情况下，可对实验现场进行录像和书面记录)。

参 考 文 献

- [1] Angelucci, M., and Giacomo De Giorgi, "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?", *American Economic Review*, 2009, 99(1), 486—508.
- [2] Ashraf, N., "Spousal Control and Intra-household Decision Making: An Experimental Study in the Philippines", *American Economic Review*, 2009, 99, 1245—1277.
- [3] Ashraf, N., D. Karlan, and W. Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines", *Quarterly Journal of Economics*, 2006, 121(2), 635—672.
- [4] Bandiera, O., I. Barankay, and I. Rasul, "Incentives for Managers and Inequality Among Workers: Evidence from a Firm Level Experiment", *Quarterly Journal of Economics*, 2007, 122, 729—774.
- [5] Banerjee, A. V., S. Cole, E. Duflo, and L. Linden, "Remedying Education: Evidence from Two Randomized Experiments in India", *Quarterly Journal of Economics*, 2007, 122(3), 1235—1264.
- [6] Becker, G., and G. Stigler, "Law Enforcement, Malfeasance, and Compensation of Enforcers", *Journal of Legal Studies*, 1974, 3(1), 1—18.
- [7] Bertrand, M., and S. Mullainathan, "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination", *American Economic Review*, 2004, 94, 991—1013.
- [8] Björkman, M., and J. Svensson, "Power to The People: Evidence from A Randomized Field Experiment on Community-based Monitoring in Uganda", *Quarterly Journal of Economics*, 2009, 124(2), 735—769.
- [9] Bohm, P., "Behaviour under Uncertainty without Preference Reversal: a Field Experiment", *Experimental Economics*, 1994, 19(2), 185—200.
- [10] Cadsby, C. B., and S. Fei, Y. Bi, "Trust, Reciprocity and Social Distance in China: an Experimental Investigation", Working Paper, 2008.
- [11] Card, D., S. DellaVigna, and U. Malmendier, "The Role of Theory in Field Experiments", *Journal of Economic Perspectives*, 2011, 25(3), 39—62.
- [12] Cameron, L., N. Erkal, L. Gangadharan, and X. Meng, "Little Emperors: Behavioral Impacts of China's One-child Policy", *Science*, 2013, 339(6122), 953—957.
- [13] Carpenter, J., and S. Erika, "Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay", *Economic Inquiry*, 2011, 49, 2, 612—630.
- [14] Carpenter, J., G. Harrison, and J. List, "Field Experiments in Economics: An Introduction", in J. Carpenter, G. W. Harrison & J. A. List (eds), *Field Experiments in Economics*. Greenwich, CT: JAI Press, Research in Experimental Economics, 2005, 10, 1—15.
- [15] Carpenter, J., S. Burks, and E. Verhoogen, "Comparing Students to Workers: the Effects of Social Framing on Behavior in Distribution Games", in J. Carpenter, G. W. Harrison & J. A. List (eds), *Field Experiments in Economics*. Greenwich, CT: JAI Press, Research in Experimental Economics, 2005, 10, 261—289.
- [16] Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick, "Optimal Defaults and Active Decisions", *Quarterly Journal of Economics*, 2009, 124(4), 1639—1674.

- [17] Chen, Y., F. Maxwell Harper, J. Konstan, and S. Li, "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens", *American Economic Review*, 2010, 100(4), 1358—1398.
- [18] Chen, K., and F. Tang, "Cultural Differences between Tibetans and Ethnic Han Chinese in Ultimatum Bargaining Experiments", *European Journal of Political Economy*, 2009, 25, 78—84.
- [19] Y. Chen, H. Ye, and D. Wang, "Theories of Social Preferences beyond Homo Economicus: A Review Based on the Experimental Economics", *Nankai Economic Studies*, 2011, 12, 63—100. (in Chinese)
- [20] Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan, "How does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR", *Quarterly Journal of Economics*, 2011, 126(4), 1593—1660.
- [21] Cohen, J., and P. Dupas, "Free Distribution or Cost-sharing? Evidence from A Randomized Malaria Prevention Experiment", *Quarterly Journal of Economics*, 2010, 125(1), 1—45.
- [22] De Mel, S., D. McKenzie, and C. Woodruff, "Returns to Capital in Microenterprises: Evidence from a Field Experiment", *Quarterly Journal of Economics*, 2008, 123(4), 1329—1372.
- [23] DellaVigna, S., J. List, and U. Malmendier, "Testing for Altruism and Social Pressure in Charitable Giving", *Quarterly Journal of Economics*, 2012, 127(1), 1—56.
- [24] Duflo, E., "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", *American Economic Review*, 2001, 91(4), 795—813.
- [25] Duflo, E., and E. Saez, "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment", *Quarterly Journal of Economics*, 2003, 118(3), 815—842.
- [26] Dufwenberg, M., and G. Harrison, "Peter Bohm: Father of Field Experiments", *Experimental Economics*, 2008, 11(3), 213—220.
- [27] Dulleck, U., J. Fookien, and Y. He, "Public Policy and Individual Labor Market Discrimination: An Artefactual Field Experiment in China", Working Paper, 2012.
- [28] Dupas, P., and J. Robinson, "Why Don't the Poor Save More? Evidence from Health Savings Experiments", *American Economic Review*, 2013, 103(4), 1138—1171.
- [29] Falk, A., and J. Heckman, "Lab Experiments Are a Major Source of Knowledge in the Social Sciences", *Science*, 2009, 326(5952), 535—538.
- [30] Fehr, E., G. Kirchsteiger, and A. Riedl, "Does Fairness Prevent Market Clearing? An Experimental Investigation", *Quarterly Journal of Economics*, 1993, 108, 437—459.
- [31] Fehr, E., and J. List, "The Hidden Costs and Returns of Incentives-Trust and Trustworthiness among CEOs", *Journal of the European Economic Association*, 2004, 2(5), 743—771.
- [32] Fisher, R. A., *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd, 1925.
- [33] Fisher, R. A., "The Arrangement of Field Experiments", *Journal of the Ministry of Agriculture*, 1926, 33, 503—513.
- [34] Gneezy, U., E. Haruvy, and H. Yafe, "The Inefficiency of Splitting the Bill: A Lesson in Institutional Design", *The Economic Journal*, 2004, 114(495), 265—280.
- [35] Gneezy, U., and J. List, "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments", *Econometrica*, 2006, 74(5), 1365—1384.

- [36] Gong, B., and C. Yang, "Gender Differences in Risk Attitudes: Field Experiments on the Matrilineal Mosuo and the Patriarchal Yi", *Journal of Economic Behavior and Organization*, 2012, 83, 59—65.
- [37] Granger, C. W. J., "Investigating Causal Relations by Econometric Models and Cross-spectral Methods", *Econometrica*, 1969, 37, 424—438.
- [38] Gurven, M., and J. Winking, "Collective Action in Action: Prosocial Behavior in and out of the Laboratory", *American Anthropologist*, 2008, 110(2), 179—90.
- [39] He, H., "The Individual and Household Trans-period Decision-making and the Subject Heterogeneity: An Analysis, Based on the Theory of Random Utility, of Experimental Economics", *Management World (Monthly)*, 2011, 12, 12—31. (in Chinese)
- [40] Harrison, G., and J. List, "Field Experiments", *Journal of Economic Literature*, 2004, 42(4), 1009—1055.
- [41] Harrison, G., M. Lau, and M. Williams, "Estimating Individual Discount Rates in Denmark: A Field Experiment", *American Economic Review*, 2002, 92(5), 1606—1617.
- [42] Heckman, J., "Econometric Causality", *International Statistical Review*, 2008, 76(1), 1—27.
- [43] Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath, "in Search of Homo Economics: Behavioral Experiments in 15 Small-scale Societies", *American Economic Review*, 2001, 91, 73—78.
- [44] Henrich, J., S. J. Heine, and A. Norenzayan, "Most People are Not WEIRD", *Nature*, 2010, 29, 466.
- [45] Jensen, R., "The (Perceived) Returns to Education and the Demand for Schooling", *Quarterly Journal of Economics*, 2010, 125(2), 515—548.
- [46] Karlan, D., "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions", *American Economic Review*, 2005, 95, 1688—1699.
- [47] Karlan, D., and J. Zinman, "Credit Elasticities in Less-developed Economies: Implications for Micro-finance", *American Economic Review*, 2008, 98(3), 1040—1068.
- [48] Katz, L. F., J. R. Kling, and J. B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment", *Quarterly Journal of Economics*, 2001, 116(2), 607—654.
- [49] Kling, J. R., S. Mullainathan, E. Shafir, L. Vermeulen, and M. Wrobel, "Comparison Friction: Experimental Evidence from Medicare Drug Plans", *Quarterly Journal of Economics*, 2012, 127(1), 199—235.
- [50] Kremer, M., and E. Miguel, "The Illusion of Sustainability", *Quarterly Journal of Economics*, 2007, 112(3), 1007—1065.
- [51] Landry, C. E., L. Andreas, J. List, M. Price, and N. Rupp, "Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment", *Quarterly Journal of Economics*, 2006, 121(2), 747—782.
- [52] Levitt, S. D., and J. A. List, "What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?", *Journal of Economic Perspectives*, 2007, 21(2), 153—174.
- [53] Levitt, S., and J. List, "Field Experiments in Economics: The Past, the Present, and the Future", *European Economic Review*, 2009, 53(1), 1—18.
- [54] List, J., "Does Market Experience Eliminate Market Anomalies?", *Quarterly Journal of Economics*, 2003, 118(1), 41—71.

- [55] List, J., "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field", *Quarterly Journal of Economics*, 2004, 119, 49—89.
- [56] List, J., "Field Experiments: A Bridge between Lab and Naturally Occurring Data", *The B. E. Journal of Economic Analysis & Policy*, 2007, 6(2 — Advances), Article 8.
- [57] List, J., and D. Lucking-Reiley, "The Effects of Seed Money and Refunds on Charitable Giving: Experimental Evidence from a University Capital Campaign", *Journal of Political Economy*, 2002, 110(8), 215—233.
- [58] List, J., and I. Rasul, "Field Experiments in Labor Economics", *Handbook of Labor Economics*, 2011, 4, Part A, 103—228.
- [59] Miguel, E., and M. Kremer, "Worms: Identifying Impacts on Education and Health in The Presence of Treatment Externalities", *Econometrica*, 2004, 72(1), 159—217.
- [60] Nagin, D. S., J. B. Rebitzer, S. Sanders, and L. J. Taylor, "Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment", *American Economic Review*, 2002, 92, 850—873.
- [61] Neyman, J., *An Outline of the Theory and Practice of Representative Method, Applied in Social Research*. Warsaw: Institute for Social Problems, 1933.
- [62] Olken, B. A., "Monitoring Corruption: Evidence from a Field Experiment in Indonesia", *Journal of Political Economy*, 2007, 115(2), 200—249.
- [63] Rosenbaum, P., and D. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, 1983, 70(1), 41—55.
- [64] Ruffle, B., and R. Sosis, "Cooperation and the In-Group-Out-Group Bias: A Field Test on Israeli Kibbutz Members and City Residents", *Journal of Economic Behavior and Organization*, 2006, 60, 147—163.
- [65] Shearer, B. S., "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment", *Review of Economic Studies*, 2004, 71, 513—534.
- [66] Smith, V. L., "An Experimental Study of Competitive Market Behavior", *Journal of Political Economy*, 1962, 70(2), 111—137.
- [67] Soetevent, A., "Anonymity in Giving in a Natural Context-A Field Experiment in Thirty Churches", *Journal of Public Economics*, 2005, 89(11—12), 2301—2323.
- [68] Sutter, M., and M. Kocher, "Trust and Trustworthiness across Different Age Groups", *Games and Economic Behavior*, 2007, 59, 364—382.
- [69] Tracer, D. P., "Selfishness and Fairness in Economic and Evolutionary Perspective: An Experimental Economic Study in Papua New Guinea", *Current Anthropology*, 2003, 44, 432—438.
- [70] Yoeli, E., M. Hoffman, D. Rand, and M. Nowak, "Powering up with Indirect Reciprocity in a Large-scale Field Experiment", *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(2), 10424—10429.

Experimental Economics Bridging to Real World: The Literature Review of Field Experiments

JUN LUO* HANG YE YEFENG CHEN

(Zhejiang University)

DINGDING WANG

(Peking University)

Abstract In this paper we summarize the research methods, core ideas, development route, classifications as defined for field experiments in economics, then we introduce some important examples which used field experiments to solve problems in economics in recent years. We also discuss the prospect of field experiment studies in China. Apart from trying to help domestic economists better understand field experiment's studies, we also provide plausible ideas and suggestions on how to carry out field experiments to bridge economics studies and the real-world in the complicated and diverse Chinese scenariocontext.

Key Words Experimental Economics, Field Experiments, Chinese Context

JEL Classification C91, C92, C93

* Corresponding Author: Jun Luo, College of Economics, Zhejiang University, 38, Zheda Road, Hangzhou, Zhejiang, 310027, China; Tel: 86-18768161869; E-mail: roger.1228@163.com.