

鲍尔斯和金蒂斯对行为和演化范式经济学的贡献^{*}

——科睿唯安“引文桂冠”经济学奖得主学术贡献评介

随学超 王立平 冯骏

摘要：鲍尔斯和金蒂斯凭借提供证据和模型以拓展对社会合作的理解，共同入选 2022 年度科睿唯安“引文桂冠”经济学奖名录。两位学者在自利合作行为模型之外，创造性地提出强互惠概念用以理解人类社会大群体中非亲缘个体之间的广泛合作，并给出其存在性证据。他们以多层次选择和基因—文化共演化理论为基础，整合生物学、经济学和社会学在内的多学科理论，坚持理论和经验主义相结合，论证群体竞争、群体间战争冲突、社会制度、社会化和惩罚在利他主义演化中的关键作用，证实利他主义可以通过自然选择得以演化。本文介绍鲍尔斯和金蒂斯的生平和学术成就，分类梳理他们的主要研究成果及学术贡献，并总结其对当前经济学研究和经济政策制定的启示。

关键词：塞缪尔·鲍尔斯 赫伯特·金蒂斯 行为和演化范式经济学 引文桂冠

塞缪尔·鲍尔斯(Samuel Bowles)为美国政治经济学家、当代激进经济学派代表人物。1939 年出生于美国康涅狄格州纽黑文市，1960 年获耶鲁大学文学学士学位，1965 年获哈佛大学经济学博士学位。博士毕业后，留在哈佛大学先后担任助理教授和副教授，1974 年开始在马萨诸塞大学阿姆赫斯特分校(UMASS Amherst)经济系担任教授至 2001 年荣休。随后，鲍尔斯被以研究复杂系统科学为使命的美国桑塔菲研究院(Santa Fe Institute, SFI)外聘为研究员，同时担任行为科学主任。2006 年因在经济理论方面的杰出贡献被全球发展与环境研究所授予里昂惕夫奖(Leontief Prize)，2020 年当选为美国艺术与科学院院士。

赫伯特·金蒂斯(Herbert Gintis)为美国政治经济学家、当代激进经济学派代表人物。1940 年出生于美国宾夕法尼亚州费城，1961 年获宾夕法尼亚大学数学学士学位，1969 年获哈佛大学经济学博士学位。博士毕业后，留在哈佛大学任教，1974 年开始转任马萨诸塞大学阿姆赫斯特分校经济系教授至 2003 年荣休。从 2001 年开始担任美国桑塔菲研究院外聘研究员，2023 年 1 月 5 日在马萨诸塞州北安普顿去世。

鲍尔斯和金蒂斯在 20 世纪 60 年代末应马丁·路德·金(Martin Luther King)博士的请求进行合作研究，就此开始了贯穿了两人整个职业生涯的合作。他们同为激进政治经济学家联盟(Union for Radical Political Economy)的创始成员，研究内容一反当时西方主流经济学的议题，早期主要关注经济不平等、资本主义国家学校教育及马克思主义经济学等问题。鲍尔斯和金蒂斯在进入桑塔菲研究院后，将研究内容扩展到行为和演化经济学领域，研究成果达到了一个新的高峰，成为桑塔菲经

^{*} 随学超、王立平、冯骏，合肥工业大学经济学院，邮政编码：230601，电子邮箱：xcsui@hfut.edu.cn。基金项目：教育部人文社会科学研究青年基金项目“农产品经纪人异质性、交易治理及其治理效应研究”(22YJC790107)；国家自然科学基金面上项目“动态博弈中有限理性局中人的前瞻性推理理论建模及行为实验”(72371090)。感谢审稿专家的宝贵意见，文责自负。

济学派的核心和领袖。2022年,鲍尔斯和金蒂斯共同凭借“提供证据和模型以扩展对不仅包括自利,还有互惠、利他主义和其他形式的社会合作等经济行为的理解”入选科睿唯安评选的“引文桂冠”经济学奖名录,本文将主要对这一学术贡献进行评介。

一、提出强互惠概念并给出其存在性证据

(一)强互惠概念的提出

1. 提出背景。人类能在大群体中与非亲属建立广泛的社会合作,正是这种能力才使智人从动物行列中脱颖而出。经济学和生物学等学科将合作定义为个体负担成本为其他个体带来收益的行为。如何理解人类社会成员之间的广泛合作是社会科学的经典难题之一。经济学家和生物学家首先试图从不违反自利假设的原则去解释。比如,互惠利他主义(reciprocal altruism)、间接互惠理论(indirect reciprocity theory)和信号成本理论(costly signaling theory)等。这些自利合作行为模型认为人们短期内在双边或多边互动中进行利他合作,是为了期待日后从受惠个体或第三方得到回报,即从长期来看仍然是自利的(Gintis, Smith & Bowles, 2001; Gintis, 2009)。上述理论尤其能够解释在双边和小型群体重复互动之中的合作行为。对于发生在亲缘关系群体中的非功利性利他合作,生物学家用亲缘利他理论(kin altruism)进行解释。

但是鲍尔斯和金蒂斯指出,自利合作行为模型不能够完全解释人类社会成员之间的广泛合作。比如,人类会在一次性互动、匿名互动、与陌生人互动以及大群体中进行利他合作,但人们从这些合作行为中很难获得预期的回报,这无法用自利合作行为模型进行解释(Gintis, 2000b, 2003a, 2013; Bowles & Gintis, 2002a)。无名氏定理(folk theorem)描述了无限重复博弈可以保证自利主体之间达成广泛的合作,但须以完全信息为条件,金蒂斯本人并不否认无名氏定理的形式正确性,但他指出人类并没有完美的信息和无限认知的能力,因此在现实生活中,自利行为者不具有足够高的远见,常表现为急功近利,导致其难以依赖重复互动在大群体中成功建立合作(Gintis, 2000a, 2004a, 2006)。同时,由于合约并不是完全的、无成本的,国家司法机构也不能够无成本地强制执行合同,因此人类社会大群体中的非亲缘个体之间的合作也无法完全利用合约进行协调(Gintis, 2004a)。

2. 强互惠。对常用于解释社会合作的利他主义,在《合作的物种:人类的互惠性及其演化》(Bowles & Gintis, 2011)一书中,鲍尔斯和金蒂斯归纳了学术界对利他主义的两种定义形式:其一,从行为的结果上进行定义,如利他主义的标准生物学定义是指,利他行为的结果能增加同群体成员的期望平均适存度(fitness),但需要自己付出净成本,这意味着行动者如果不做出这一行为,就可以增加自己的适存度。其二,从激励利他行为的偏好上进行定义,利他主义者具有这样一种偏好,会给有利于他人的行为结果赋予正的价值,这一利他偏好会激励行动者在某些时候承担个人成本做出利他行为,而这一偏好的属性可以是“自虑”(self-regarding)的,也可以是“他虑”(other-regarding)的。

在《强互惠与人类社会性》(Gintis, 2000b)一文中,金蒂斯提出了一种他虑的利他主义——“强互惠”(strong reciprocity),指人们合作不仅仅是出于自利的原因,还出于对他人福利的真诚关心,以及固有的遵守合作规范的倾向。人们会赋予符合道德的行为本身以正的价值,出于同样的原因,人们也会不惜花费个人成本去惩罚那些违反合作规范的人(即使违反不是针对自己)。金蒂斯在“互惠”一词前面加“强”字,是为了将其与传统自利性质的互惠即“弱互惠”(weak reciprocity)进行区别(Gintis, 2004c)。弱互惠只有在有回报预期的前提下才会发生,因此是一种自虑行为;而强互惠行为并不是基于有回报预期才会发生,强互惠者能够从这种利他行为本身获得满足感(Gintis et al, 2008; Gintis, 2013)。金蒂斯特别指出,强互惠为无回报预期的情况下进行合作提供了一种动机,群体中存在一定数量的强互惠者,他们不但实施合作行为,即实施“利他合作”(altruistic cooperation),而且不惜花费成本惩罚违反合作规范的群体成员,即实施“利他惩罚”(altruistic punishment),从而抑制群体中的背

叛、卸责和搭便车行为，这对于维持大群体中非亲缘个体之间的合作秩序意义重大(Gintis, 2017a)。

鲍尔斯和金蒂斯在后续一系列文献中对强互惠行为和偏好的内涵进行了充分的论述，他们指出强互惠大多数表现为一种有条件的利他主义，而不是无条件仁慈。比如，当他人表现出合作时，强互惠者也倾向于合作；当他人违反规范时（即使违反不是针对自己），哪怕需要个人承担成本也要对其进行惩罚，即“以德报德，以牙还牙，惩恶扬善”(Fehr & Gintis, 2007; Gintis, 2013)。金蒂斯还特别指出，强互惠行为既不是适应不良(maladaptation position)，也不是头脑混乱或者盲目情绪化，而是一种符合强互惠者他虑偏好的理性行为(Gintis, 2011)。从行为的结果上看，强互惠者个人付出的成本不能从合作收益中得到完全补偿，会导致净损失，因此强互惠是符合“从行为结果定义”的利他主义；从偏好的性质来看，驱动强互惠行为的偏好是人们对他人福利的关心和维护道德规范的愿望，因此强互惠是一种他虑或者完全道德的利他主义。

3. 强互惠行为的激励机制。一个显而易见的问题是，在不能从外界获得完全的物质收益补偿的情形下，人们为什么愿意花费个人成本进行合作以及惩罚违反合作规范的个体？金蒂斯认为，强互惠行为是受到个体他虑或道德的偏好驱动的，与人类演化得到的骄傲、羞耻、内疚、愤怒等社会情感紧密相关，强互惠者之所以进行利他合作和利他惩罚，是因为这样的行为会激起满足和骄傲的情感，如果不这么做，就会感到羞耻和内疚(Gintis et al, 2008)。这一社会情感的激励机制得到了广泛的行为实验和其他证据的支持，研究人员发现人们能够从惩罚违反规范者的行为中获得精神满足(Gintis, 2017b)。

(二) 强互惠存在的经验证据

1. 古人类社会和现代社会生活的经验观察。强互惠偏好和行为是人类社会进化的晚期产物还是在古人类社会就已经存在？在整理大量考古学、遗传学、民族志和历史学文献资料的基础上，鲍尔斯和金蒂斯指出，在晚更新世和全新世早期的古人类社会，就已经存在广泛的社会合作行为，如在族群内部存在集体狩猎、食物分享以及在群体防卫中进行合作等实践，且发展了一系列的利他惩罚行为以维持合作秩序(Bowles & Gintis, 2011; Gintis, 2017b)。同时，他们依据文献资料确凿地指出，祖先人类可能生活在流动性高、遗传亲缘程度低的大群体中，存在与陌生人互动的可能，仅凭亲缘利他和自利互惠难以维持广泛的社会合作。因此，金蒂斯指出，强互惠动机可能在古人类社会就已存在，否则无法解释古人类社会就已经存在的广泛的社会合作行为(Gintis, 2013)。

金蒂斯进一步指出，人类在现代社会生活情境中也会表现出普遍的强互惠行为，比如匿名向慈善机构捐款、参与政治运动、支持面向穷人的转移支付政策、遵守社会规范、自愿在战争中战斗和牺牲以及受害者对犯罪者的“自我毁灭性报复”等(Gintis, 2008, 2015)。对个人而言，人类在这些行为中付出的成本会显著超出可能获得的收益，这些行为难以用自利或弱互惠理论来解释，金蒂斯认为这些行为具有强互惠行为的关键属性。

在《互惠、自利与福利国家》(Bowles & Gintis, 2000a)一文中，鲍尔斯和金蒂斯极其细致地讨论了选民对福利国家的支持和反对这一现象。在过去，经济学家普遍认为选民对福利国家的支持是因为自利动机，即认为支持收入再分配是风险厌恶型公民偏爱的一种保险形式。然而，经验证据显示，随着一国收入不平等程度的加剧，选民对收入再分配的支持程度反而下降，这是自私选民理论(selfish voter theory)不能很好解释的现象。与传统的经济学解释相反，他们在该文中指出，选民之所以支持福利国家，是因为其符合人类根深蒂固的强互惠偏好，选民即使在没有预期回报情形下，也会选择支持福利国家，但支持福利国家的选民也并非是无条件的利他主义者。针对近几十年来发达国家选民对福利国家的支持率下降的情况，两位学者指出，这是因为社会福利计划未能充分兑现对公平的承诺——选民认为社会福利计划助长了不道德的生活方式和价值观（比如懒惰、非婚生子等），这表明支持福利国家的动机是一种有条件的利他主义(Bowles & Gintis, 2000a)。鲍尔斯和金蒂斯利用强互惠动机解释选民对福利国家的支持和反对是能够逻辑自洽的，这也进一步为强

互惠的存在提供了经验证据支持。

2. 来自行为实验的证据。古人类社会和现代生活情境难以为一次性和匿名互动提供空间,因此人们表现出的行为不能排除不是一种自虑行为。而行为实验方法可以控制非匿名性与重复互动的影 响,使我们能够对人类的偏好和行为做出相当具体的推论(Fehr & Gintis,2007)。金蒂斯及其合作者在一系列文章中通过分析经典文献中的博弈实验结果,给出了强互惠存在的实验经济学证据,这些实验包括最后通牒博弈(ultimatum game)、公共产品博弈(public goods game)和礼物交换博弈(gift exchange game)等。

在《强互惠与人类社会性》(Gintis,2000b)一文中,金蒂斯指出,强互惠可以很好地解释“最后通牒博弈悖论”。在来自各种文化背景的相当多的最后通牒博弈实验中,提议者往往给回应者提供慷慨的分配金额,总量40%~50%的分配提议是最普遍的,回应者往往会拒绝低于总量25%的分配提议。金蒂斯指出,提议者的行为或许可以用自利动机去解释,但回应者的行为是不能够用自利动机来解释的,也许惩罚不公平的提议本身就是最有可能的动机。此外,在由计算机生成的出价(而不是由人提议)中,低报价被拒绝的可能性显著降低。这一事实表明,哪怕自己付出代价也要拒绝低分配提议是为了惩罚违反公平规范的人,而不是简单地拒绝不利的出价结果。

金蒂斯进一步指出,强互惠也能恰如其分地解释标准的多期公共产品博弈实验中的“期数效应”(round effect)(Gintis,2000b,2011)。在标准的多期公共产品博弈实验中,刚开始时,玩家捐献的水平很高,但在以后的轮次中,捐献逐渐衰退,即存在所谓的期数效应。金蒂斯认为,这是因为一部分玩家是强互惠者,表现出条件性合作(conditional cooperation)倾向所导致的(Gintis,2011)。强互惠者在实验伊始就向公共账户捐献反映了玩家的利他合作倾向,但当发现有玩家不捐献时,他们就会动用拒绝或减少捐献的方式来惩罚搭便车者(玩家事后回忆的报告),因此导致了标准公共产品博弈实验中捐献逐期衰退。当允许被试花费个人成本向搭便车者施加惩罚时,实验结果显示,对搭便车者的惩罚行为普遍存在,而且在最后几轮博弈中,惩罚者也没有减少(Fehr & Gächter,2000)。金蒂斯认为,这表明即使不能改变搭便车者或潜在搭便车者的后续行为,惩罚者也会不惜代价对搭便车者进行惩罚,即惩罚不是出于自利的动机,而仅仅因他们违反了合作规范(Gintis,2000b)。特别值得一提的是,允许对搭便车行为进行惩罚能够维系较高的捐献水平。

在《对人类利他行为的解释》(Gintis,Bowles,Boyd & Fehr,2003)一文中,金蒂斯及其合作者指出,经典的礼物交换博弈的结果也揭示了强互惠的存在性。在礼物交换博弈中,给定雇主—雇员的互动是一次性事件,这意味着,即使雇员违反契约约定的努力水平,也不会受到惩罚。根据自利假设预期,不管接受怎样的契约,雇员均会选择零成本努力。然而,现实实验结果却显示,雇主提供的工资越慷慨,雇员提供的努力程度就越高。金蒂斯等认为,该结果表明雇员具有强互惠倾向,即雇员会付出一定的努力以回应雇主的慷慨。在扩展的礼物交换博弈中,当允许雇主支付成本对雇员进行奖惩时,在68%的情况下,雇主惩罚了没有履行契约的雇员,而在70%的情况下,雇主奖励了超额完成契约的雇员。金蒂斯等进一步认为,扩展实验的结果表明雇主也具有强互惠倾向,即使惩罚并不能为自己提供未来的预期收益,雇主也会不惜花费个人成本进行“奖好惩坏”。

在《合作的物种:人类的互惠性及其演化》(Bowles & Gintis,2011)一书中,鲍尔斯和金蒂斯回顾了许多经典的博弈实验文献,这些实验在具有不同经济文化背景的地区的实验室或自然现场(natural field)进行了成百上千次重复,实验的外生参数设置虽然有些许差异,但实验结果对自利“经济人”假设构成了系统性背离。所幸的是,鲍尔斯和金蒂斯提出的强互惠动机能很好地解释这些行为实验的结果。

二、利他主义演化的解释

诸如强互惠形式的利他主义,具有个人代价高昂但对社会有益的属性,这必然使利他主义者在

与自利者的竞争中呈现劣势,其如何能通过自然选择得以进化呢?这一问题便是著名的“人类亲社会之谜”。在经济学家和生物学家看来,只有对利他主义的起源及其演化做出合理的解释,才能倾向于相信从社会生活和行为实验的经验证据中归纳出的利他偏好和行为。经济学家通常用自然选择必然产生自私有机体这一观点证明传统主流经济学自利假设的正确性,鲍尔斯和金蒂斯试图挑战这一观点,即自利假设不仅与来自社会生活和行为实验的经验证据不一致,而且也不是自然选择的必然结果,诸如强互惠等形式的利他主义也可以通过自然选择得以演化。本文接下来将重点评介两位学者在这一领域的理论与实证工作,以厘清其理论的内在逻辑和具体的实证发现。

(一) 群体间竞争主导利他主义演化

不可否认,国家、公司等群体会面临竞争压力和优胜劣汰,而自然选择可以在群体层面发挥作用这一观点,早在达尔文提出进化论之后的另一部著作中就已经被提及。在《强互惠与人类社会性》(Gintis, 2000b)一文中,金蒂斯基于群体选择(group selection)的思想,以多人公共物品博弈作为分析框架构建了一个理论模型,模型显示,强互惠可以通过群体选择在自利互惠中产生。在该文中,金蒂斯指出,拥有利他主义者的群体能够维持群体合作,增加群体收益,因此在竞争中会战胜其他的自利群体。特别是当人类群体在面临灭绝威胁时(如遇到战争、瘟疫或者饥荒),基于自利互惠的合作将面临崩溃,此时只有强互惠者的存在才能维持合作,拥有一定强互惠者的群体在这种情况下可以更好地生存,使得强互惠者可以和自利者共同演化。

但是生物学家对群体选择理论持谨慎态度,原因有以下两个方面:一方面,认为群体选择理论无法解决“内部颠覆”(subversion from within)问题。例如,由于在群体内部自私个体相较利他主义个体有适存度优势,因此无法抵制自私的基因在群体内部的蔓延,也无法抵御自私基因的外部侵入。另一方面,利他主义群体要战胜其他群体,进而在自然选择中胜出,须以两个群体中利他主义者的比例存在实质性的差异为条件,但群体之间的迁徙、等位基因的不断重组和融合往往会限制群体基因构成的差异性。如何回应这些质疑,是学术界利用群体选择理论解释利他主义演化面临的一个重要挑战。

鲍尔斯认为,社会中的实体总是表现为不同层面的群体和个体,因此,在利用群体选择理论分析利他主义的演化时,并不排斥个体选择(individual selection)理论,倾向于一种多层次选择(multilevel selection)理论的解释,即利他主义的演化取决于群体间和群体内选择的相对大小(Bowles, 2006)。在《群体竞争、繁衍均衡与人类利他主义的演化》(Bowles, 2006)一文中,鲍尔斯指出,虽然群内选择倾向于淘汰利他主义者,但是群间选择更偏爱利他主义者,如果群体间遗传差异足够大、群体间竞争足够激烈,群体选择就可以抵消利他主义者面对的群内选择压力。鲍尔斯在该文中建立了一个包含群体间遗传差异、群体间竞争强度和群体社会结构等关键因素在内的理论模型,然后利用计算机仿真估计了利他主义得以演化的参数条件,发现激烈的群体间竞争有利于利他主义的演化。鲍尔斯还在该文中特别指出,考虑到群体分裂、群体灭绝和选择性配型(selective assortment)的影响,重新估计的人类祖先族群间的遗传差异值已经足够大,在这样的参数空间内,群体间竞争可以解释利他主义的演化。

(二) 群体间战争冲突与利他主义演化

对人类利他主义演化的解释,关键不在于群体间竞争是否会影响利他主义演化,而是人类祖先族群之间是否存在足够激烈的群体间竞争事实。对此,鲍尔斯指出,人类祖先族群间存在频繁战争冲突这一事实加剧了群体间竞争,强化了群体选择效应,致使代价高昂的利他主义也能扩散,并形象地将战争冲突称为利他主义的“助产士”(Bowles, 2009, 2012)。在《狩猎—采集者族群间的战争是否影响了人类社会行为的进化?》(Bowles, 2009)一文中,鲍尔斯构建了群体间竞争演化影响模型,利用计算机仿真得到的结果显示,当设置较高的战争冲突致死率参数值时,即使群体内利他主义者数

量比较低,也能保证该群体在群体间竞争中胜出,并可以支持具有极高成本的利他主义的演化。基于民族志和考古学的资料印证了,在晚更新世和全新世早期族群之间,存在着频繁且致命的战争冲突,这也进一步支持了群体选择在人类利他主义的演化解释中扮演的重要角色。

但是上文的解释又引发了一个问题,即为什么人类祖先族群之间会存在频繁战争冲突?鲍尔斯和金蒂斯指出,在晚更新世和全新世早期,异常的气候波动常导致自然灾害和周期性资源短缺,数量庞大的人类狩猎—采集族群为了争夺资源,会陷于频繁且异常致命的群体间冲突。此外,鲍尔斯等还给出了人类祖先族群之间战争冲突的一种内生解释。鲍尔斯和金蒂斯提出了一种在人类社会生活和行为实验中普遍观察到的狭隘主义(parochialism)概念,即对群体内成员表现出高度的利他主义行为,但敌视群体外的成员(Bowles & Gintis, 2004a)。鲍尔斯认为,狭隘主义是人类祖先族群间战争冲突的产物,同时,狭隘主义也会导致群体间的暴力冲突,并且有助于在冲突中胜出(Bowles, 2008a)。鲍尔斯与其合作者基于晚更新世和全新世早期的数据,利用计算机仿真得到的结果显示,经过数千代自然选择后出现的种群(population)往往要么主要由宽容的非利他主义者构成且很少有战争,要么主要由狭隘主义者构成但存在频繁的群体间战争冲突(Choi & Bowles, 2007)。这一结果也表明,狭隘主义与战争冲突是一种共演化关系,这也为人类祖先族群间为什么存在频繁的战争冲突提供了一种内生解释。

晚更新世和全新世早期是人类祖先进化的关键时期,人类许多的认知特征,包括与利他主义有关的特征都是在那一时期进化的。鲍尔斯和金蒂斯努力从考古学、气候学、遗传学、民族志和历史学等研究中汲取关于这一时期的古人类所处的社会和自然环境的信息,以尽可能精确地勾勒出古人类社会生活的图景,并据此对模型仿真参数进行校准,保证模型参数设置在经验上是合理的,以此增强各种理论解释的严谨性和可信性。两位学者的研究并不仅仅依赖理论模型提出相关观点,而是花费相当大的篇幅来证明,模型的假设和结论与我们了解到的古人类社会生活相一致。

(三) 社会制度与利他主义的共生演化

从狩猎—采集社会到民族国家,在人类社会发展的各个阶段中,都存在一些既普遍又独特的社会制度,社会生物学将其称为“演化普遍物”(evolutionary universals),把这些制度的涌现归因于文化的群体选择(cultural group selection),而基于文化传递(cultural transmission)的社会制度对“基因传递”的利他主义性状在自然选择下的演化至关重要(Bowles, 2004)。在《个人行为与社会制度的共同演化》(Bowles, Choi & Hopfensitz, 2003)一文中,鲍尔斯等基于多层次选择和基因—文化共演化(gene-culture coevolution)理论,详细地论证了一些特殊的社会制度通过减轻利他主义者面临的群内选择压力,以支持利他主义在大群体中得到稳定演化的机制。

在晚更新世和全新世早期的人类狩猎—采集社会普遍存在两种制度,即繁衍均等化(reproductive leveling)和分片化(segmentation)。繁衍均等化是指那些能降低群体内成员适存度或减少物质福利差异的制度,如资源和信息共享、协商一致的决策、集体限制占优者的独断行为、一夫一妻制等。而群体内分片化是指,群体内同类型个体更有可能互动的倾向,如利他者更有可能与利他者互动,自利者也更有可能与自利者互动,而不是随机配对。在《个人行为与社会制度的共同演化》(Bowles, Choi & Hopfensitz, 2003)一文中,鲍尔斯等指出繁衍均等化和分片化可以降低利他主义个体相对同群体其他成员的适存度劣势,减轻利他主义者面临的群内选择压力,从而可以支持利他主义在大群体中的演化。鲍尔斯等在该文中基于基因—文化共演化理论,构建了一个社会制度和个人行为共同演化的模型,采用计算机仿真演绎模型的结果显示:在设置繁衍均等化和分片化制度可以自由演化时,群体层面的选择就会支持利他主义行为与两种社会制度共同演化;在限制两种社会制度演化的情形下,只有在群体间冲突非常频繁、群体规模较小和迁徙率较低的情况下,群体层面的选择才能支持利他主义行为的演化。而晚更新世和全新世早期的社会和自然环境有可能恰好落在模型设置的仿真参数

空间范围内,因此其所构建的制度和个人行为共演化模型也许至少可以提供关于这一关键时期个体利他行为与繁衍均等化、分片化等制度演化的部分解释。

鲍尔斯等在该文中还特别指出,基于文化传递的社会制度更能够维持群体间的差异,因此可以强化利他主义演化的群间选择效应(Bowles, Choi & Hopfensitz, 2003)。对于基因传递的个体性状,在群体间迁徙的影响下,等位基因的不断重组和融合往往会稀释群体间基因构成的差异性,群体选择在其进化中就很难起作用。制度是群体层面的属性,群体边界能够维持制度结构中群体差异的稳定性,而制度的延承是一种基于社会学习(social learning)的文化传递过程。因此,即使存在群体间移民,只要新移民通过社会学习的方式习得迁入地的制度规范,表现出与迁入地成员一致的行为性状,就可以保证群体间个体行为性状差异的稳定性。

(四) 社会化与利他主义演化

1. 社会化与利他规范的演化。人类社会存在非常普遍的规范内化(internalization of norms)现象,而规范内化主要通过基于文化传递的社会化(socialization)实现,具体是指社会成员从父母(纵向传递)、社会化机构(斜向传递)和同伴(横向传递)习得本群体社会规范和社会价值观的过程。金蒂斯认为,当社会规范得到内化,人们就会把遵守规范本身当做值得追求的偏好,而人们遵守规范主要由内部制裁(internal sanctions)强制执行,比如感到羞愧和内疚等,而不是依赖于外部奖惩(Gintis, 2006)。利他规范在社会中广泛存在,社会学学科普遍认同使用规范内化来解释人类的亲社会行为。鲍尔斯和金蒂斯对利他主义演化的解释也没有无视这一社会学学科的观点。社会规范是文化的重要内容,文化演化理论认为人类文化与基因包含的信息在代际间的传递具有相似性,都要受到突变和选择的影响。由于利他规范对内化者具有收益劣势,使内化者承担了较大的群体内选择压力,因此利用规范内化理论来解释利他主义的演化需要解决的一个关键问题是如何解释利他规范的演化。

对此,在《破解亲社会之谜》(Gintis, 2003a)一文中,金蒂斯提出了一个文化演化模型,分析了利他规范在文化传递的影响下是如何演化的,并从文化均衡方面解释了群体内利他主义持续存在的原因。模型假定父代向子代传递的是自己的社会规范(无论是利他规范还是利己规范),社会化机构(如世俗、宗教仪式、学校和传媒等机构等)只传递利他规范。同时还假定,利他型社会成员会模仿有更高适应度的自利型同伴的社会规范,受此影响会以一定概率转化为自利型个体。模型分析结果显示,当不存在基于社会化机构的斜向传递时,利他主义就会被利己行为所驱逐;当存在基于社会化机构的斜向传递时,利他主义可以在文化均衡中持续存在,表明家庭外的社会化对于利他主义的演化是必要的。同时还发现,利他行为的成本越高,利他规范得以涌现的条件就越苛刻,群体内存在的少数内化互惠规范的成员可以在文化均衡中维持高水平的合作。

相较基因传递和个体学习,金蒂斯指出,文化传递是一种人类特有的、高等的信息获取方式,这依赖于人类拥有的认知、语言及其他生理能力,以及规范内化能力(Gintis, 2003a)。智人是一种进化而来的高度社会化的生物物种,至少近万年以来,人类文明的演化动力主要来源于文化传递,远超过纯粹自然进化的选择力量(Bowles, 2008a)。如果说群体竞争对于解释古人类社会利他主义的初始出现和扩散具有重要意义,那么利用文化演化模型解释现代人类社会异乎寻常的合作就有其充分的必要性。金蒂斯构建的文化演化模型综合了生物学、经济学和社会学在解释人类行为方面的核心观点,有助于准确及创新性地理解利他主义的演化。例如,社会学学科对规范内化的一个批判是认为规范内化过于强调社会对个体的影响,认为人完全屈从于经由社会化过程内化的价值观和规范体系,即“过度社会化”(oversocialized)。而金蒂斯在其构建的文化演化模型中引入了模仿者动态效应,可以对抗个体的过度社会化,指出人的行为不仅受到内化的价值观和规范体系的影响,还会受到经济动机的影响。这一理论框架综合了经济学和社会学的核心观点,也更贴近于人类在现实中的行

为表现。

2. 基因—文化共演化与规范内化能力。家庭和社会化机构通过基于文化传递的社会化过程将偏好传递给子代,但须以人类的规范内化能力为前提。那么,人类是否具有这种规范内化能力呢?如果人类拥有这一能力,但是内化一个会降低自身适应性的规范,既违反了经济学自利“经济人”假定,也违反生物学进化思想,规范内化能力又怎样通过自然选择呢?

对此,金蒂斯指出,来自社会的观察证实了人类普遍存在内化规范能力,而且认为这一能力是由基因控制的,因此可以在代际间进行稳定传递,而“反社会者”的存在通常是因为内化规范能力的弱化或缺失,并认为这是具有遗传倾向的(Gintis,2003b)。来自社会学的观点认为,内化规范能力具有重大的适应性价值,在《合作的物种—人类的互惠性及其演化》(Bowles & Gintis,2011)一书中,鲍尔斯和金蒂斯对此进行了阐述:其一,成功的社会中存在大量能增强个体适存度的规范,如未雨绸缪、良好的个人卫生、良好的工作习惯和情绪控制等;其二,内化规范能力缓解了个体在每一种情境中计算成本和收益的压力,并减小了计算失误的概率;其三,拥有内化规范的能力可使得人类能在新形势下进行快速的文化适应,而纯粹的基因适应从接受指令到发挥效用需要花费更长的时间,智人发展时期的重要特征之一就是存在快速的文化进化和高度可变的环境。

在《利他主义作为搭便车者的解释:基因—文化共演化及规范内化》(Gintis,2003b)一文中,金蒂斯通过构建基因—文化共演化的多个模型证明,由于存在能增强个体适存度的规范,拥有内化规范能力的个体具有适应性优势。因此,内化规范能力能够得到演化,而降低适存度的利他规范可以通过“搭便车”得到内化。在利他规范的适存度成本小于斜向传播强度时,利他规范的表型就会变得演化可行。金蒂斯同样坚持理论和经验主义相结合,在同时考虑基因选择、习得行为以及模仿者动态效应的情形下,利用计算机对上述基于主体(agent-based)的理论模型进行仿真。结果显示,仿真通常很快就能达到稳定状态,文化传递和规范内化能力能够协同演化,在稳定时,利他主义和斜向传播强度均可以维持在一个高水平。计算机仿真的结果还显示,即使设置较高的群体间迁徙率,也没有破坏利他均衡。鲍尔斯和金蒂斯指出,这是因为多数效应都发生在文化层面而不是基因层面,这能够抑制群体间的差异被移民稀释(Bowles & Gintis,2011)。

(五)利他惩罚的演化

对违反合作规范的群体成员实施惩罚可以诱导社会合作,减轻合作者的群内选择压力,长期来看,可以引导个体内化那些导向合作行为的规范。因此,利他惩罚可以支持利他合作的演化。但利他惩罚本身就是利他主义的一种形式,因此具有适存度劣势的利他惩罚的演化也受到相当多的理论和实证关注。大多数文献对利他惩罚演化的解释通常基于人类具有以低成本实施严厉惩罚的能力的假设,此时通过维持合作带来的群体收益可以补偿惩罚者的适应性劣势。这在一定程度上是合理的,因为人类具有制造可发射的武器和实施道德攻击的独特能力,可以确保以低成本实施严厉惩罚(Gintis,2000b)。可是这一假设日益受到学术界质疑,惩罚不仅需要惩罚者付出成本,以及降低受罚者的适存度,同时还可能波及利他合作者。另外,在现实中惩罚还可能引致报复,这也会使惩罚变得异常昂贵,这些质疑在一定程度上给解释利他惩罚演化带来挑战。针对这些质疑,鲍尔斯和金蒂斯通过对现实中的惩罚行为本身的细致深入分析,构造模型展示了惩罚如何作为一种社会协调活动得到内生演化。

1. 利他惩罚和利他合作的非对称性机制。在《利他惩罚的演化》(Boyd, Gintis, Bowles & Richerson,2003)一文中,金蒂斯及其合作者认为,利他惩罚和利他合作之间具有重要的非对称性,利他合作者相对于搭便车者的收益劣势与群体中搭便车行为的频率无关,而利他惩罚者的成本劣势会随着搭便车者人数的下降而下降。在一个利他惩罚者普遍的群体中,大多数成员将参与合作,此时惩罚很少发生,利他惩罚者几乎不承担成本,那么其面临的群内选择压力会非常微弱。该模型基于多

人公共品重复博弈展开,模型的计算机仿真结果显示,当没有惩罚时,群体选择只有在小群体规模情况下才能维持高频率的合作;当允许惩罚存在时,群体选择可以在大群体规模的参数值下维持利他惩罚和利他合作的稳定演化。由于利他合作和利他惩罚的非对称性机制,群体选择在维持利他惩罚方面比利他合作更有效。

2. 协调惩罚。在《对背叛者的协调惩罚维持了合作并且能够得到扩散》(Boyd, Gintis & Bowles, 2010)一文中,金蒂斯及其合作者指出,民族志的证据显示,惩罚并非是无条件的或者没有协调的个体自发行为,而是通常呈现为协调惩罚(coordinated punishment)形式,即惩罚是要通过惩罚者之间的交流进行协调的,且通常是集体行为。基于此,金蒂斯等提出了一个协调惩罚演化模型。首先,将惩罚视为协调惩罚,惩罚是否发生取决于倾向参与惩罚的成员数量,当愿意惩罚的个体很少时,惩罚不会发生,“惩罚者”只需要承担发出惩罚意愿信号的成本,避免了昂贵的惩罚成本。其次,当有足够多的惩罚者时,惩罚发生,但可以共同分担惩罚的成本。因此,协调惩罚可以降低个体惩罚成本或增加惩罚力度。模型以典型的人类祖先狩猎—采集社会的互动结构为基础,利用计算机仿真的结果显示,惩罚搭便车者的总成本会随着惩罚者数量的增加而下降,惩罚者罕见时也可以扩散,而当惩罚者很常见时,会增加群体的平均适存度。

3. “二阶搭便车”困境与强互惠的演化。如前文所述,利他惩罚可能具有昂贵的成本,所以与纯粹合作者相比,强互惠者具有适存度劣势,对于个人而言,选择始终合作但拒绝惩罚可能是一个占优策略。如果个体选择遵守规范,但是拒绝实施惩罚或分担惩罚的成本,则可以坐享惩罚带来的好处。如此就可能产生“二阶搭便车”(second-order free ride)困境,即一旦所有人都选择不惩罚,群体内全部由拒绝惩罚的合作者组成时,自私者又会侵入该群体并逐步取代合作者。

在《强互惠的演化:异质人群中的合作》(Bowles & Gintis, 2004b)一文中,鲍尔斯和金蒂斯为强互惠者如何在“二阶搭便车”困境中得以演化提供了一个可信的解释。他们首先假定存在三种类型的个体:自私者——竭力逃避合作责任的卸责者;合作者——始终合作但拒绝惩罚卸责者;强互惠者——既合作又惩罚卸责者,对卸责者实施的惩罚形式为排斥(ostracism)和疏远(shunning)。该文通过建立7个基于主体的方程以模拟晚更新世时期的狩猎—采集种群合作演化过程。计算机仿真结果显示,即使初始状态全是自私者,由突变产生的较小数量的强互惠者也可以侵入自私者人群中,且强互惠者能在与自私者、合作者的生存竞争中得到稳定演化。仿真结果还显示,只有当一个族群演化出强互惠者后,才能建立起稳定的合作秩序,并达到演化均衡。模型中强互惠者的惩罚行为能够降低自私者的适存度,从而减轻合作者、强互惠者面对的群内选择压力。完全由合作者构成的族群是不稳定的,最终会被自私者统治。而对于完全由自私者组成的族群,由于缺乏合作机制维持的适应性优势,最终将走向灭绝。

三、对经济学研究和经济政策制定的启示

新古典主流经济学最基本的“经济人”假设与现实中的丰富人性之间存在着较大的差距,基于经济人假设和完全且无成本的契约建立的新古典主流经济学体系自20世纪20年代以来受到了严峻的挑战。鲍尔斯和金蒂斯运用行为和演化范式来研究经济学,不仅突破了新古典主流经济学研究范式的桎梏,而且其研究成果在经济政策制定中也具有截然不同的启示。

(一)对经济学研究的启示

与基于外生自利假定的新古典经济学研究范式相比,鲍尔斯和金蒂斯所主张的行为和演化范式的经济学研究具有显著的差异性特征。在《回顾瓦尔拉斯经济学》(Bowles & Gintis, 2000b)、《内生偏好:市场和其他经济制度的文化后果》(Bowles, 1998)等文献中,鲍尔斯和金蒂斯严厉质疑了新古典主流经济学的外生自利偏好假设,并指出经济学研究应摆脱这一不合时宜的假设,经济行为的基

础需要重新表述。鲍尔斯和金蒂斯认为,在特定情形下,人类行为决策模型应将他虑偏好等异质性偏好作为目标纳入效用函数,而不是把它们视为约束条件(Bowles & Gintis, 2000b)。在此基础上,金蒂斯提出了理性行动者(beliefs, preferences and constraints, BPC)模型,该模型可以接受偏好的异质性,能够对人类千差万别的行为进行建模,同时模型还能将决策者的认知心理因素包括进来(Gintis, 2007a; Gintis, 2009)。BPC模型为解释人类行为的异质性提供了一个统一框架,特别是对研究那些行为和实验经济学文献中列出的偏好不一致性和行为非理性的“异象”具有重要意义。

长期以来,新古典主流经济学将对偏好的解释归于心理学、生物学和社会学的研究领域,而经济学主要关注固定偏好下的资源配置效率问题。这一观点在一定程度上简化了经济学研究的任务,但也限制了经济科学的解释力。鲍尔斯指出,偏好并不总是一成不变的,而是深刻地受到生活环境、社会规范和文化传统的复杂影响,如果不考虑偏好的内生性,经济学既不能准确预测,也不能准确评估新政策可能产生的后果(Bowles, 1998)。偏好内生理论对未来经济学研究的拓展方向具有重要启示意义。首先,基于个体偏好和制度的内生互动关系,研究偏好和制度的形成和演化成为拓展经济学研究范围的一个有价值方向(Bowles & Gintis, 2000b)。其次,偏好的内生性意味着在经济建模时需要考虑偏好与制度间的反馈效应,制度的影响效应可能超出了以往研究所强调的范围,这一反馈效应可能改变市场的结果,应当引起研究重视(Bowles, 1998)。再次,对偏好的解释涉及心理学、生物学和社会学学科,经济学需要与其他行为学科进行更紧密的结合,才能更好地理解人类行为和偏好的形成和演化(Gintis, 2004b, 2007b)。

(二)对经济政策制定的启示

在新古典主流经济学中,无论主张依靠市场还是政府进行资源配置,其解决问题的背后都有一个共同的隐含机制,即利用人们的自利动机以实现社会公共目标。鲍尔斯及其合作者在《社会偏好与公共经济学:社会偏好依赖于激励的机制设计》(Bowles & Hwang, 2008)、《为自利公民设计的政策可能破坏“道德价值观”:来自经济实验的证据》(Bowles, 2008b)等文献中,对基于新古典主流经济学的“经济人”政策范式给予了批评。鲍尔斯及其合作者的研究显示,传统的基于自利“经济人”的政策和制度设计,即通过改变目标活动的经济成本和收益以实现公共目标,可能会削弱个人遵守社会规范和为公共利益做贡献的道德动机,导致激励效果降低,而且从长期看会破坏人们的道德价值观,鼓励个人的行为更加自私(Bowles, 2008)。因此,对政策制定者来说,考虑人类利他主义、互惠和遵守道德规范的倾向等亲社会偏好尤为重要,良好的制度和政策设计不仅要利用物质利益激励来实现社会目标,同时也要能增强道德情感动机对这一目标的贡献(Bowles & Hwang, 2008)。需要指出的是,古典主义者在政策制定时并不是不考虑道德情感动机,而是认为物质利益激励和道德情感对行为的影响是可分离的,物质利益激励和道德情感激励是相互独立的,这被称为分离性假设(assumption of separability)。但是,鲍尔斯及其合作者给出的证据显示,这一假设通常是错误的(Bowles & Hwang, 2008; Bowles, 2008b)。

新古典主流经济学主张应以市场作为现代社会的主要治理方案,对于偶尔的市场失灵则采用政府干预,而社区治理(community governance)被视为未开化的中世纪时期的残余。社区的内涵是指以多种方式进行直接、频繁互动的社会群体。但过去50年的实践表明,市场失灵逐渐成为一种常态而不是例外,并且政府没有足够的力量来纠正所有的市场失灵。例如,在现代经济社会,交易互动、商品和服务性质均变得更加复杂,使得人们很难签订完备程度较高的契约,也增加了法律处理的成本,市场治理和政府治理的优势降低(Bowles & Gintis, 2002b)。在这样的背景下,在《社会资本与社区治理》(Bowles & Gintis, 2002b)一文中,鲍尔斯和金蒂斯重提利用社区治理体系来解决典型的市场和政府失灵。社区治理依赖于分散的私人信息,这些信息通常是国家、雇主、银行和其他大型正式组织无法获取的,社区能够根据成员符合或偏离社会规范的情况对其实施奖惩。与国家相

比,社区能更有效地利用社会规范来实施激励。两位学者并不否认社区治理的缺陷,但同时也指出,这并不能成为否定社区治理的理由,应通过制度设计,使社区治理、市场治理和政府治理互为补充,而不是相互替代(Bowles & Gintis,2002b)。

四、结语

鲍尔斯和金蒂斯作为激进主义经济学派的代表人物,在学术生涯的开始就一反主流经济学的议题。在进入桑塔菲研究院后,他们将研究内容扩展到行为和演化范式经济学领域,主要关注违反主流经济学假设的利他行为和偏好。两位学者在自利合作行为模型之外,创造性地提出强互惠概念,将其界定为他虑的或完全道德的行为和偏好,用以理解人类社会大群体中非亲缘个体之间的合作,并通过整理分析大量有关晚更新世和全新世早期古人类社会生活的文献资料以及经典的行为实验文献,给出强互惠偏好和行为的存在性证据。

鲍尔斯和金蒂斯对“人类亲社会之谜”的一个核心观点是,个人代价高昂的利他主义是适应性的,是自然选择的结果,显著区别于经济学家的传统观点,即自私个体是自然选择的必然结果。两位学者对这一观点的论证借鉴了来自进化生物学的多层次选择和基因-文化共演化理论:一方面,种群中的总选择可以分解为群体层面和个体层面上的选择;另一方面,在人类进化中,基因与文化之间存在紧密的互动,人类的偏好是基因影响文化演化、文化影响基因进化的动态过程的结果。两位学者系统性地研究了人类利他主义演化的多个关键机制。首先,晚更新世和全新世早期面临着一个异常的气候波动环境,在争夺有限资源、应对自然灾害的竞争中,群体间竞争使得自然选择倾向于利他主义者,而人类祖先族群之间频繁战争冲突也可以在群体层面上强化有利于利他主义者的自然选择。其次,在群体内部,一些既普遍又独特的社会制度,如繁衍均等化和分片化等减少了利他主义者的群内选择压力,从而可以支持利他主义在大群体中的稳定演化。再次,社会化是人类利他主义演化的又一关键因素,这依赖于人类的内化规范能力。两位学者指出,基于文化传递的社会规范与基于“基因传递”的人类规范内化能力可以协同演化,而降低个体适存度的利他规范在社会化的作用下可以战胜利他者所承受的适存度劣势,通过规范内化得以在代际间进行稳定传递。最后,人类社会在漫长的发展中演化出的惩罚行为能够降低利他主义的群内选择压力,利他惩罚的存在可以支持利他主义在极为苛刻的条件下得到演化。而且,即使代价昂贵的利他惩罚,也能作为一种社会协调活动得到内生演化,群体选择在维持利他惩罚上要比利他合作更有效。两位学者的研究方法也很值得关注,例如,利用跨学科的分析建模去解决一个有高度争议的科学问题,基于主体的建模和计算机仿真可以有针对性地处理行为演化的复杂性,理论和经验事实的深度融合增强了解释的严谨性和可信性。

鲍尔斯和金蒂斯运用行为和演化范式来研究经济学,不仅突破了新古典主流经济学研究范式的桎梏,而且其研究成果在经济政策制定中也具有截然不同的启示,对于解决当今经济社会发展的现实难题具有重要的启示意义。当前我国经济社会已经进入一个新的发展阶段,经济社会治理面临着诸多新课题。行为和演化范式经济学彻底抛弃了新古典经济学不切实际的“经济人”假设,关注现实生活中人们如何互动、市场如何运作。这一理论体系可以包容人类行为的异质性、偏好的内生性以及经济系统的复杂性,两位学者的研究对当前经济学研究的拓展方向具有重要的启示意义。2008年全球金融危机以来,西方发达国家在应对贫富分化、社会不平等和金融市场周期性动荡等经济社会难题方面陷入困境,基于新古典主流经济学的“经济人”政策范式被认为是上述问题的根源。在此背景下,基于偏好内生性、偏好异质性以及他虑偏好的制度和政策设计可作为破解当前经济社会难题的一个有力工具。

参考文献:

Bowles, S. & H. Gintis(2000a), “Reciprocity, self-interest, and the welfare state”, *Nordic Journal of Political*

Economy 26(2):33—53.

- Bowles, S. & H. Gintis(2000b), “Walrasian economics in retrospect”, *Quarterly Journal of Economics* 115(4):1411—1439.
- Bowles, S. & H. Gintis(2002a), “Homo reciprocans”, *Nature* 415(6868):125—128.
- Bowles, S. & H. Gintis(2002b), “Social capital and community governance”, *Economic Journal* 112(483):F419—F436.
- Bowles, S. & H. Gintis(2004a), “Persistent parochialism: Trust and exclusion in ethnic networks”, *Journal of Economic Behavior & Organization* 55(1):1—23.
- Bowles, S. & H. Gintis(2004b), “The evolution of strong reciprocity: Cooperation in heterogeneous populations”, *Theoretical Population Biology* 65(1):17—28.
- Bowles, S. & S.-H. Hwang(2008), “Social preferences and public economics: Mechanism design when social preferences depend on incentives”, *Journal of Public Economics* 92(8—9):1811—1820.
- Bowles, S. & H. Gintis(2011), *A Cooperative Species: Human Reciprocity and Its Evolution*, Princeton University Press. 中译本题名:《合作的物种:人类的互惠性及其演化》(鲍尔斯和金迪斯著),浙江大学出版社 2015 年版。
- Bowles, S. (1998), “Endogenous preferences: The cultural consequences of markets and other economic institutions”, *Journal of Economic Literature* 36(1):75—111.
- Bowles, S. (2004), *Microeconomics: Behavior, Institutions, and Evolution*, Princeton University Press. 中译本题名:《微观经济学:行为,制度和演化》(鲍尔斯著),中国人民大学出版社 2006 年版。
- Bowles, S. (2006), “Group competition, reproductive leveling, and the evolution of human altruism”, *Science* 314(5805):1569—1572.
- Bowles, S. (2008a), “Conflict: Altruism’s midwife”, *Nature* 456(7220):326—327.
- Bowles, S. (2008b), “Policies designed for self-interested citizens may undermine ‘the moral sentiments’: Evidence from economic experiments”, *Science* 320(5883):1605—1609.
- Bowles, S. (2009), “Did warfare among ancestral hunter-gatherers affect the evolution of human social behaviors?”, *Science* 324(5932):1293—1298.
- Bowles, S. (2012), “Warriors, levelers, and the role of conflict in human social evolution”, *Science* 336(6083):876—879.
- Bowles, S., J. K. Choi & A. Hopfensitz(2003), “The co-evolution of individual behaviors and social institutions”, *Journal of Theoretical Biology* 223(2):135—147.
- Boyd, R., H. Gintis & S. Bowles(2010), “Coordinated punishment of defectors sustains cooperation and can proliferate when rare”, *Science* 328(5978):617—620.
- Boyd, R., H. Gintis, S. Bowles & P. J. Richerson(2003), “The evolution of altruistic punishment”, *PNAS* 100(6):3531—3535.
- Choi, J.-K. & S. Bowles(2007), “The coevolution of parochial altruism and war”, *Science* 318(5850):636—640.
- Fehr, E. & H. Gintis(2007), “Human motivation and social cooperation: Experimental and analytical foundations”, *Annual Review of Sociology* 33(1):43—64.
- Fehr, E. & S. Gächter(2000), “Cooperation and punishment in public goods experiments”, *American Economic Review* 90(4):980—994.
- Gintis, H. (2000a), *Game Theory Evolving*, Princeton University Press. 中译本题名:《演化博弈论》(金迪斯著),中国人民大学出版社 2015 年版。
- Gintis, H. (2000b), “Strong reciprocity and human sociality”, *Journal of Theoretical Biology* 206(2):169—179.
- Gintis, H. (2003a), “Solving the puzzle of prosociality”, *Rationality and Society* 15(2):155—187.
- Gintis, H. (2003b), “The hitchhiker’s guide to altruism: Gene-culture coevolution, and the internalization of norms”, *Journal of Theoretical Biology* 220(4):407—418.
- Gintis, H. (2004a), “Modeling cooperation among self-interested agents: A critique”, *Journal of Socio-Economics* 33(6):695—714.
- Gintis, H. (2004b), “Towards the unity of the human behavioral sciences”, *Politics, Philosophy & Economics* 3(1):37—57.

- Gintis, H. (2004c), “Economic interests: Do strangers cooperate when they have to work together?”, *Nature* 431 (7006):245–246.
- Gintis, H. (2006), “Behavioral ethics meets natural justice”, *Politics, Philosophy & Economics* 5(1):5–32.
- Gintis, H. (2007a), “A framework for the unification of the behavioral sciences”, *Behavioral and Brain Sciences* 30 (1):1–16.
- Gintis, H. (2007b), “Unifying the behavioral sciences II”, *Behavioral and Brain Sciences* 30(1):45–53.
- Gintis, H. (2008), “Punishment and cooperation”, *Science* 319(5868):1345–1346.
- Gintis, H. (2009), *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*, Princeton University Press. 中译本题名:《理性的边界:博弈论与各门行为科学的统一》(金蒂斯著),格致出版社2011年版。
- Gintis, H. (2011), “Gene-culture coevolution and the nature of human sociality”, *Philosophical Transactions of the Royal Society B – Biological Sciences* 366(1566):878–888.
- Gintis, H. (2013), “An implausible model and evolutionary explanation of the revenge motive”, *Behavioral and Brain Sciences* 36(1):21–22.
- Gintis, H. (2015), “Altruists together”, *Nature* 517(7536):550–551.
- Gintis, H. (2017a), “How humans cooperate and punish”, *Nature Human Behaviour* 1(9):626–627.
- Gintis, H. (2017b), *Individuality and Entanglement: The Moral and Material Basis of Social Life*, Princeton University Press. 中译本题名:《个体性与纠缠:社会生活的道德与物质基础》(金蒂斯著),格致出版社2021年版。
- Gintis, H., E. A. Smith & S. Bowles(2001), “Costly signaling and cooperation”, *Journal of Theoretical Biology* 213(1):103–119.
- Gintis, H., S. Bowles, R. Boyd & E. Fehr(2003), “Explaining altruistic behavior in humans”, *Evolution and Human Behavior* 24(3):153–172.
- Gintis, H., J. Henrich, S. Bowles, R. Boyd & E. Fehr(2008), “Strong reciprocity and the roots of human morality”, *Social Justice Research* 21(2):241–253.

Contributions of Bowles and Gintis to Behavioral and Evolutionary Paradigm Economics

SUI Xuechao WANG Liping FENG Jun
(Hefei University of Technology, Hefei, China)

Abstract: Bowles and Gintis were named 2022’s Citation Laureates in Economics by Clarivate for providing evidence and models that broaden our understanding of social cooperation. Apart from the cooperation model based on self-interested behavior, Bowles and Gintis creatively proposed strong reciprocity to explain the human propensity to cooperate extensively with nonrelatives, and interpret the experimental evidence as showing the existence of strong reciprocity. Based on the theory of multi-level selection and gene-culture coevolution, they integrated multidisciplinary approaches including economics, sociology and biology, and insisted on the combination of theory and empiricism to demonstrate the role of group competition, inter-group warfare, social institutions, socialization and punishment in the evolution of altruism, and confirmed that the costly altruism is the result of natural selection. This article introduces the lives and academic achievements of Bowles and Gintis, classifies and combs their research achievements and academic contributions, and summarizes the enlightenment to current economic research and policy-making.

Keywords: Samuel Bowles; Herbert Gintis; Behavioral and Evolutionary Paradigm Economics; Citation Laureate

(责任编辑:李仁贵)

(校对:刘洪愧)