

协调性惩罚与人类合作的演化^{*}

韦 倩 孙瑞琪 姜树广 叶 航

内容提要：虽然惩罚可以维持合作的演化，却给惩罚者带来自然选择的适应性劣势，从而形成新的演化之谜。人类在漫长的发展中逐渐演进各种巧妙的协调性惩罚行为，这无疑会降低惩罚成本或增加惩罚力度，当它们改变惩罚行为的成本劣势，使其达到成功演化的阈值范围时，反对惩罚的选择性压力就消失，从而可以支持利他性惩罚在大规模群体中演化。本文不再遵循惩罚者同质性假设，从而能够讨论各种协调性惩罚对群体合作秩序演进的影响。本文构建动态演化模型，模拟了人类祖先的一个典型狩猎采集社会，在其他参数不变的情况下，设置了各种协调性惩罚的环境，并对比分析了它们对合作演化的影响。仿真结果表明，在规模较大的情况下，非协调性惩罚很难维持群体合作的演化，而协调性惩罚则会明显提高群体合作水平，特别是在一致行动和权威惩罚的情景中，群体可以实现高水平的合作。

关键词：协调性惩罚 合作 演化 模拟

一、引言

在十余万年之前，就影响力和重要性而言，人类的先祖与水母、蟑螂、老虎等其他物种别无二致，但是现在人类已经成为地球的统治者（Harari, 2014）。人类能够从毫不起眼的人猿走到这一步，不是由于身体、智力等方面的个体优势，而是因为更善于合作。不过，合作行为具有利他性，合作者自己付出成本却使其他非亲缘成员获益，从而不具有演化上的优势（Boyd et al., 2003; Gintis, 2000）。对于人类合作的演化之谜，理论解释主要包括互惠利他理论^①（Trivers, 1971; Axelrod & Hamilton, 1981）、间接互惠理论^②（Nowak & Sigmund, 1998; Wedekind & Milinski, 2000）、有成本的信号理论^③（Gintis et al., 2001）、利他性惩罚理论^④（Gächter et al., 2008; Boyd et al., 2003）、文化群体选择理论^⑤（Boyd & Richerson, 1985）、网络互惠理论^⑥（Nowak & May, 1992）等。

^{*} 韦倩，山东大学经济研究院，邮政编码：250100，电子信箱：weiqian1979@163.com；孙瑞琪，山东大学经济研究院博士研究生，邮政编码：250100，电子信箱：ruiqi_sun1990@126.com；姜树广、叶航（通讯作者），浙江财经大学经济学院、浙江财经大学经济行为与决策研究中心，邮政编码：310018，电子信箱：jsg123000@sina.com, yehang@china.com。本文系国家自然科学基金面上项目（71873076）、教育部人文社会科学研究规划基金项目（17YJA790076）、山东省自然科学基金面上项目（ZR2016GM25）的阶段性成果，同时得到了国家“万人计划”青年拔尖人才（W03070176）、山东省“泰山学者”青年专家（TSQN20161008）人才项目的资助。作者感谢匿名审稿专家的宝贵意见，但文责自负。

① 互惠利他理论认为，如果未来重复交易的机会足够高，那么在双边交易中自然选择会偏爱合作行为。

② 间接互惠理论认为，个体之所以选择合作行为是期望以此获得良好声誉，从而今后能得到第三方的奖励，而不是仅来自受惠个体的回报。

③ 有成本的信号理论认为，人们可以通过合作行为传递其作为潜在伙伴的信号，这使他们具有结盟优势，从而导致合作在无亲缘关系的社会群体中成功演化。

④ 利他性惩罚理论认为，由于背叛可能会遭受严厉的惩罚，使得合作者具有比背叛者更高的适应性，因此惩罚可以维持群体合作的演化。

⑤ 文化群体选择理论认为，个体的合作行为有利于整个种群，因此合作行为会跟随群体的留存而得以繁衍。

⑥ 网络互惠理论认为，社会交往存在空间结构或网络差异，这种差异会影响群体的生物动态性，合作者通过空间聚集来避免被背叛者剥夺，从而可以促进合作的演化。

利他性惩罚理论无疑是其中最具影响力的解释之一。但是,同合作一样,惩罚也是有成本的,这会形成一个新的演化之谜,即减少个人适应性的惩罚行为如何得以成功演化?针对此问题,大多数文献主要基于补偿的角度进行解释,即惩罚行为带来的收益可以抵补其行为本身给惩罚者带来的适应性劣势,比如基于群体选择(Gintis, 2000; Bowles & Gintis, 2004)、社会顺从(Henrich & Boyd, 2001)、社会排斥(Sasaki & Uchida, 2013; Panchanathan & Boyd, 2004)、社会交往的空间结构(Helbing et al., 2010; Szolnoki & Perc, 2013)、多重惩罚(Boyd & Richerson, 1992)、变异(Sethi & Somanathan, 1996)等视角。上述文献对于理解人类合作之谜提供了非常有价值的解决思路,但其缺点也显而易见,一是它们假设惩罚成本不能太高,即惩罚对背叛者造成的伤害要大于惩罚者付出的成本。但令人沮丧的是,这条假设在现实中并不总是成立。惩罚者并不一定是背叛者的对手,并且背叛者经常采取报复行为,这使得惩罚成本有时非常高(韦倩, 2009)。很明显,在惩罚成本较高的情况下,仅仅依靠补偿机制的解释就很难让人信服。二是缺乏对惩罚行为本身细致深入的分析。^①我们认为,具有高级认知能力的人类在面对背叛行为时,既不会像其他物种那样转身离去,也不会简单、机械地对背叛者进行惩罚,而是会利用各种机制协调对付背叛者,因此对惩罚行为细致深入的分析或许可以更清晰地理解人类合作演化的真相。

本文尝试脱离补偿机制的视角,将注意力放在惩罚行为本身上,这使得本文分析更加贴近现实。我们认为,惩罚行为更多是在惩罚者之间协调进行的,这些协调无疑会降低惩罚成本或增加惩罚力度,当它们改变惩罚行为的成本劣势,使其达到成功演化的阈值范围时,反对惩罚的选择性压力就消失,从而可以支持利他性惩罚在大规模群体中演化。要检验某一事件或因素对结果的影响,一般使用两种方法:一是实验,二是模拟。实验方法使得变化的不同要素被隔离观察,从而可以发现每个后果的充分原因。近年来,实验室实验和田野实验的方法被广泛应用于许多研究中,比如周业安和宋紫峰(2008)。但是,本文并未采用实验的方法,原因在于,我们认为人天生就是社会性生物,惩罚意愿与协调技巧深植于意识结构中,很多情况下都是神经系统自动完成的,这使得很难通过控制实验条件获得一个作为参照系的纯粹自发惩罚(即非协调性惩罚)的环境。因此,本文选择了模拟的方法,即通过选定某一系统的关键特征与参数,利用模型复现实际系统发生的本质过程。

本文通过构建动态演化模型,模拟了人类祖先的一个典型狩猎——采集社会,研究惩罚的协调性对惩罚以及合作演化的影响,并对比分析了不同的协调机制及其作用。本文的贡献主要体现在三点:一是虽然认为惩罚对于合作的成功演化起至关重要的作用,但与已有大部分文献不同,我们不认为惩罚行为是简单、机械地进行的,而是存在着协调。这一方面更加符合现实,另一方面即使在单个惩罚者惩罚成本较高的情况下,协调性惩罚也能维持群体合作的稳定演化。二是将协调性惩罚进行了细分,将其分为一致行动以及代理惩罚两类,又进一步将代理惩罚分为雇佣惩罚和权威惩罚两种情况,然后对比研究这三类协调性惩罚对于群体合作演进的不同影响。三是在模型设定细节上,与Jaffe & Zaballa(2010)、Boyd et al. (2010)等文献不同,我们不再假设惩罚者是同质的,而是假定他们的惩罚能力禀赋异质,且服从正态分布,这一方面更符合实际,另一方面也使讨论和比较各类协调性惩罚对群体合作的影响成为可能。

本文以下部分的结构安排为:第二部分主要阐述协调性惩罚的相关概念、分类等,并讨论其起源;第三部分设置了一个惩罚行为不存在任何协调的情景,并构建模拟仿真的基本分析框架,探讨

^① 也有少数例外,比如Boyd et al. (2003)察觉到了惩罚与合作的本质差异。他们认为,合作的成本劣势与群体中背叛者的人数无关,但惩罚的成本劣势却会随背叛者人数的下降而下降,当惩罚行为普遍时,群内反对惩罚的选择性压力相当微弱,正是这种非对称性的机制支持惩罚在大规模群体中演化。与该文献不同,本文把研究重点放在协调性惩罚对惩罚成本的影响上。

惩罚对该群体合作及演进的影响,从而为后文分析建立一个比较基准;第四部分主要探讨惩罚存在协调性情景下群体的合作及演进,并比较各种协调性惩罚的效果;第五部分为本文结论。

二、惩罚的协调性及其起源

所谓协调性惩罚,是指惩罚者相互配合实施的惩罚行为。它既可以表现为惩罚者通过默契或协商等方式达成的一致行动,还可以表现为惩罚者通过委托代理实施的惩罚行为,以及社会中各种惩罚制度安排等。其实,协调性惩罚更符合人类祖先的现实。惩罚并非总是无条件的或者是没有协调的个人自发行为,而是常常通过小道传闻或与其他惩罚者的沟通而得以协调进行(Wiessner, 2005; Boehm, 1993)。人类学的许多证据也表明,个体很少独自对违背社会规范者进行惩罚,惩罚通常以集体的形式发生,惩罚本身就传递了同伴谴责的信号。Gintis(2000)更是明确指出,通过使用协调、秘密行动和致命武器等手段,少数目的攻击者也可以通过低适应性成本来击败最强大的单个敌人。可见,协调性惩罚可以降低惩罚成本或增加惩罚力度,使得惩罚得以繁衍并增加群体收益。

根据成本降低机制的不同,将协调性惩罚分为以下两种:一是一致行动,^①即惩罚者共同实施惩罚行为,即使他们的作用有主次之分。一致行动主要依靠分摊成本、规模效应等机制降低成本或增加惩罚力度。二是代理惩罚,即惩罚者委托其他人或组织代替自己进行惩罚。与委托方相比,代理方的惩罚能力一般会更强,正是惩罚在不同个体之间的委托代理行为降低了惩罚成本。一致行动与代理惩罚虽然都存在协调,但它们最本质的区别在于:前者中所有有意愿惩罚的个体都参与了惩罚,而后者中一部分具有惩罚意愿的个体退化为普通合作者,他们委托那些进化为专职惩罚者的个体来惩罚,本身并不进行惩罚。至于现实中,为何会选择这种协调性惩罚而不是其他类型,有偶然性因素,但主要原因是群体成员关于协调成本与惩罚效果之间的利弊权衡。这种权衡既可能是理性选择的结果也可能是自然进化的结果。

但是,在协调性惩罚的起源问题上,我们更倾向于认为,与其说是理性选择的结果,不如说是来自于灵长类甚至更远古祖先的遗传。人天生就是社会性生物,对人类近亲大猩猩的行为研究表明,人类及其灵长类祖先从没作为隔离的个体存在过,在人类出现之前,社会交往就已经是其行为的一部分(De Waal, 1989; Wrangham & Peterson, 1996)。这说明霍布斯等人提出的自然状态根本不曾存在过。霍布斯认为,人类一开始各行其是,仅在发展中较迟阶段才进入社会,即通过理性推理发现社会合作是达到各自目标的最佳方法。经济学也沿袭了这一传统,通常把行为或规范的形成看成是一个理性的过程,比如社会、利他行为等都可以被归结为追求自身利益的结果。但真实的情况应该是,百万年前狩猎采集社会的环境塑造了人类的惩罚行为及其相关技巧(当然包括协调性惩罚),并将其编码于意识中,从而遗传下来,不必在每一代身上重新培养。其实,不止惩罚,整个自然界的秩序都产生于盲目和非理性的生物演进和自然选择过程(Dawkins, 1986),生物界所呈现的高度秩序不是上帝或某位造物主之赐,而是来自较低等生命的互动(福山, 2015)。

在协调性惩罚的繁衍发展过程中,语言和宗教起到了非常重要的促进作用。语言可以传递信息从而促进人类在惩罚行为上的协调。Harari(2014)认为,智人在大约7万年前发展出新的语言技能,这使他们能够传播流言蜚语、“八卦”数小时之久,从而可以保护群体成员免遭欺诈和被占便宜,发展出更紧密、更复杂的合作形式,而猿类、尼安德特人以及更早期的智人则没有这种技能。另外,语言的发展还可以促进大脑的发育,并使抽象和理论成为可能。许多学者认为,人类同黑猩猩一支分道扬镳后,新近发展出来的大脑主要用于发挥语言功能(Pugh, 1977)。而黑猩猩虽然有足

^① 也可以把这种协调性惩罚称为结盟,但名称的不同不会改变事情的本质。

够的记忆和沟通技巧,但是没有语言,从而无法达到更高层次的合作。除了语言外,宗教也是凝聚社会的源泉。宗教可以让人达成一致意见和具有相同的偏好,从而有利于惩罚协调性的发生。特别随着群体规模的扩大,社会制裁机制的缺失使得搭便车等机会主义愈发明显,集体行动开始瓦解,而宗教可以缓解此类问题。没有宗教,人类社会很难超越族群的层次。目前,考古学家还没有发现没有宗教的原始社会,尼安德特人和其他原始人类群体也可能有宗教信仰(Wade,2009)。

三、非协调性惩罚情景下的合作及其演进

为了设置一个比较基准,我们构建了一个不存在协调性惩罚的场景,即假定惩罚行为都是独立和自发进行的,不存在任何协调行为。虽然与现实不符,但是,这样做可以对比分析协调性惩罚对群体合作演化的影响。

1. 模型框架

假设存在一个进行公共物品博弈的规模为 n 的人群,^①该人群由合作者、背叛者、惩罚者三种类型的个体组成,除了惩罚者的惩罚能力有所不同之外,同一类型中的个体无差异。假定每次博弈时,群体中合作者的比例为 x ,惩罚者的比例为 y ,则背叛者的比例为 $1-x-y$ 。合作者个人支付成本 c 产生了一个可以被群体所有成员均分的总收益 b ;背叛者既不支付成本也不产生收益;惩罚者会先合作,然后再对群体中的每一个背叛者付出 p 单位的成本进行惩罚。由于惩罚能力是不同的,假设惩罚者 i 付出 1 单位成本可以给背叛者造成 $\alpha_i > 0$ 倍的代价,其中, $i = 1, 2, 3, \dots, ny$, 则每位背叛者付出的代价为 $k = p \sum_{i=1}^{ny} \alpha_i$ 。那么,单个合作者的期望收益为 $P_c = b(x+y) - c$, 单个惩罚者的收益为 $P_p = b(x+y) - c - (1-x-y)np$, 单个背叛者的收益为 $P_b = b(x+y) - k$ 。由于惩罚需要成本,导致惩罚者的收益小于合作者的收益,因此,理性的惩罚者不会对背叛者进行任何惩罚。而当群体中不存在惩罚者时,背叛者的收益就变为 $b(x+y)$, 大于合作者的收益。这样,背叛者会最终取代合作者,该群体的合作秩序会崩溃,陷入社会困境。

假设该群体的演化服从莫兰过程(Moran process)。在任意时间步,随机死亡一个原有个体,且同时产生一个新个体,从而群体数量始终保持不变。死亡个体的类型取决于该类型个体的比例,新个体的类型则取决于该类型个体的演化适应度(Nowak,2006)。虽然生物的演化适应性与其收益成正比,但二者并不完全一致,考虑到适应性不能为负,适应性方程采取特劳森函数(Traulsen et al.,2008)形式,这样,合作者的适应性为 $\pi_c = e^{\omega P_c}$, 背叛者的适应性为 $\pi_b = e^{\omega P_b}$, 惩罚者的适应性为 $\pi_p = e^{\omega P_p}$, 其中 ω 为自然选择强度。另外,假设基因会以一定的概率 μ 发生突变,这样,新产生个体的类型实际上会受到适应度和突变的双重影响。

2. 仿真模拟

本文的仿真平台基于 Netlogo 构建而成。^② 许多仿真均是依托该平台进行的,比如董志强和李伟成(2019)等。在本仿真中,合作者、背叛者和惩罚者的数量,合作的成本 c 和收益 b , 以及惩罚者付出的惩罚成本 p , 均可以任意设定,从而群体规模也是可变的,这样可以通过改变其中某一个或某几个变量来探究其对合作的影响。仿真过程中的每一时间步(tick)均由四个步骤组成:

Step1: 捐献过程。群体中所有成员参与公共物品博弈,合作者和惩罚者自行承担合作成本,收益则贡献给整个群体用于集体分配,在此阶段,背叛者能够通过一定的方式进行卸责。

Step2: 分配过程。分配遵循平均分配原则,即所有成员共享合作成果。由于背叛者并未付出

① 公共物品博弈反映了人类早期狩猎—采集社会的特征。

② 资料来源: <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

合作成本,在参与收益分配后,背叛者的实际收益将大于合作者和惩罚者的收益。

Step3:惩罚过程。惩罚者可以发现背叛行为,并能在付出一定成本后对其进行惩罚,惩罚能力的数值在惩罚者出生之时被赋予。

Step4:死亡和繁殖过程。该步骤主要基于莫兰过程,随机死亡一个原有个体,且同时伴随着一个新个体的产生,死亡个体的类型取决于该类型个体的比例,新个体类型则取决于该类型个体的演化适应度。考虑到现实中存在基因突变,我们将突变引入到繁殖过程。

在每一时间步完成后,自动进入下一个时间步中,如此循环往复,直到达到初始设置的循环次数为止。各参数的取值情况如下:①群体规模 n :Boehm(2007)分析了154个搜食者社会的特征,发现这种高度合作的游牧型多家庭群体的规模尽管存在季节性变化,但通常包括大约20—30个个体。Marlowe(2005)根据民族志记录中与人类祖先生活环境类似的175个狩猎—采集群体得出群体的平均规模为37人。^①因此,本仿真设定 n 的取值范围为 $n \in [6, 50]$ 。②合作成本 c 与合作收益 b :根据常见的公共物品博弈实验,选取 $c = 1, b = 2$ 或 $b = 4$,这样,合作者个人支付1单位成本可以产生2或4单位的被所有群体成员均分的总收益。③惩罚成本 p 与惩罚力度 α_i :根据常见公共物品博弈,令惩罚成本 $p = 0.15$ 或 $p = 0.1$;假设惩罚力度 α_i 服从正态分布 $\alpha_i \sim N(4, 4)$,且 $1 \leq \alpha_i \leq 10, \alpha_i \geq 1$ 是因为人类具有高超的制造工具和狩猎的能力,比如投掷抛射物等(Darlington, 1975; Fifer, 1987),从而可以在一定程度上以低成本对受罚者施加严厉处罚,平均来看,惩罚者施加1单位惩罚成本可以给背叛者带来4单位的伤害。④各类型人群的初始分布:对群体中各类型人群初始分布没有设置严格约束,这样, $0 \leq x \leq 1, 0 \leq y \leq 1$,因为初始分布对模拟结果的影响不显著。⑤基因突变率 μ : $\mu \in \{0.0001, 0.001, 0.01\}$ 。⑥自然选择强度 $\omega = 0.5$ 。

3. 仿真结果分析

在上述参数取值范围内,进行了数百次仿真,结果表现出很大程度上近似性。图1显示了一个代表性仿真的结果,各参数的取值情况如下: $n = 30, x = 10, y = 10, c = 1, b = 4, p = 0.15, \alpha_i \sim N(4, 4)$ 且 $1 \leq \alpha_i \leq 10, \mu = 0.001, \omega = 0.5$ 。从图1可以看出,在非协调性惩罚情景下的100万代仿真中,背叛者在人群中占据主导地位(合作者的平均比例为1.72%;背叛者的平均比例为98.02%;惩罚者的平均比例为0.26%)。可见,在不存在其他补偿机制的情况下,单纯依靠惩罚并不能维持群体的合作。即使突变产生合作者或惩罚能力很强的惩罚者,也会在短暂存活后趋于灭绝,因此背叛者占优的局面很难逆转。这表明,在一个进行公共物品博弈的较大规模群体中,即使惩罚者可以实施比较严厉的惩罚,也很难维持群体的合作。

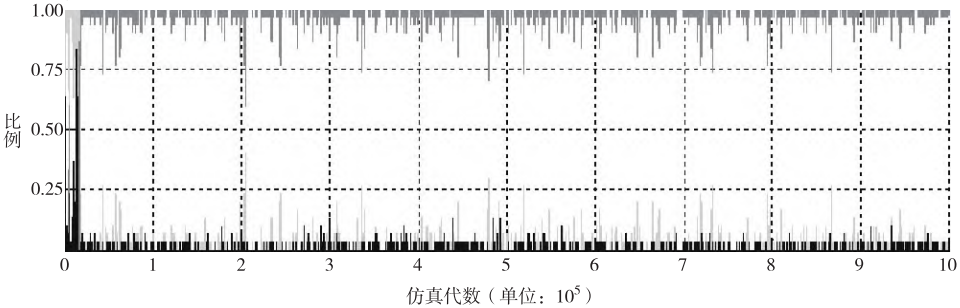


图1 非协调性惩罚情形下的人群分布比例

注:浅灰色表示合作者,深灰色表示背叛者,黑色表示惩罚者。下同。

① Bowles & Gintis(2011)认为,典型规模要比平均规模更能反映群体规模状况,他们推测 Marlowe(2005)样本中典型群体规模为77,要比平均规模大两倍以上。虽然缺乏更精确的关于早期人类祖先群体规模的数据,但是,对于规模超过50人以上的群体,我们仍然进行了仿真,其结果表现出与本文所述相似的特征。

为了更清楚地观察群体中各类型成员比例的变化情况,专门截取了一些关键时期的仿真数据进行详细分析,图 2 与图 3 分别显示了第 1—1000 代与第 16201—17200 代的仿真情形。如图 2 所示,在初始阶段,由于惩罚者的比例较多($y = 1/3$),背叛者逐渐减少,合作者和惩罚者逐渐增加,在大约 250 代之后,背叛者逐步消失殆尽。此时,由于合作者相对惩罚者具有适应性的优势,从而合作者可以入侵惩罚者,从大约 620 代之后,惩罚者逐渐减少,在大约 860 代之后,该群体全部是合作者。当群体中全部是合作者时,由于缺乏惩罚者的监督惩罚,背叛者可以完成对合作者的逆袭,这个过程大约发生在 16201—17200 代之间,如图 3 所示,大约在第 17081 代之后,该群体基本上全部被背叛者充斥。在非协调性惩罚情景下,由于没有相应的降低惩罚成本或增加惩罚强度的机制,在今后长达接近 100 万代的过程中,即使突变出惩罚者与合作者,他们再也没能侵入背叛者策略,如图 1 所示。这个结果与多数文献也一致,文献研究表明,当群体规模很大且成员之间互动交往很少时,惩罚的未来收益不能被内化,自发的同辈惩罚(peer punishment)效果就不太明显,只有在群体的规模足够小(Boyd et al.,2010)或个体有不合作的选择权利(Fowler,2005;Hauert et al.,2007)等限制条件下,自发惩罚才有效。

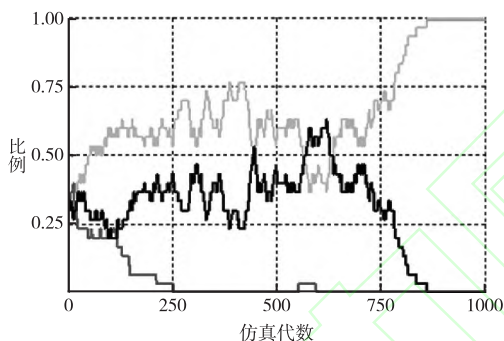


图 2 0—1000 代的仿真

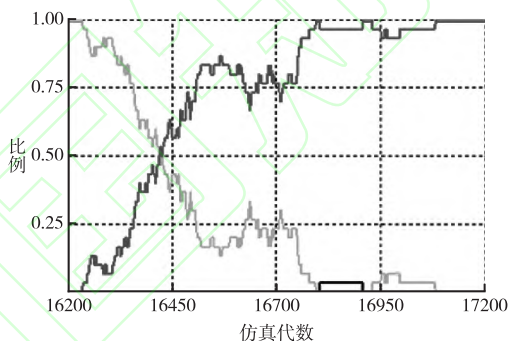


图 3 16201—17200 代的仿真

四、协调性惩罚情景下的合作及其演进

其实,非协调性惩罚并不符合人类社会的真实情况。由于具有语言、认知等与众不同的能力,人类演化发展出了各种各样的协调性惩罚行为,正是它们维持了人类社会的合作秩序。我们将协调性惩罚分为两种:一是一致行动,二是代理惩罚。代理惩罚又可细分为两种:一是雇佣惩罚,二是权威惩罚。本部分主要研究这三类协调性惩罚如何影响了群体合作的演进。在本文的分析框架中,除了惩罚方式本身的差异,其他参数均不发生变动。这样,演化结果的差异在很大程度上便可确定为不同惩罚方式的影响,通过与非协调性惩罚机制这个参照基准相比,可以发现协调性惩罚对人类合作演化的重要影响。

1. 一致行动

为了方便对比,沿用上一部分分析框架,仍然假定每个惩罚者对每一个背叛者付出的惩罚成本为 p ,但是,与上一部分的独立惩罚不同,惩罚者之间可以选择采取一致行动。假定具有惩罚意愿的个体需要先进行沟通(既包括协商等正式形式,也包括流言、八卦等非正式形式),才能进行一致行动。假定沟通也是需要成本的,其比例为 δ ,这样,惩罚成本就分为直接惩罚 $p(1 - \delta)$ 和协调成本 $p\delta$ 两部分。必须要强调的是, $p\delta$ 为惩罚者投入在协调或沟通中的成本,而不像交易成本一样,是社会或群体中实际存在的成本。假定直接惩罚给惩罚者带来的伤害强度与非协调性惩罚相同,为 $p(1 - \delta) \sum_{i=1}^{n_y} \alpha_i$;协调成本虽然不能对背叛者造成伤害,但是它可以放大惩罚者的惩罚能力,假定其放大幅度与协调成本所占比例 δ 、惩罚者人数 n_y 成正比。这样,惩罚者采取一致行动给每位背叛

者带来的伤害为 $k = ny\delta(1-\delta)p \sum_{i=1}^{ny} \alpha_i$ 。对于一个理性的惩罚者来说,他的最优选择是在惩罚成本 p 一定的情况下给背叛者造成的伤害 k 最大。那么,如果一致行动的惩罚给背叛者带来的代价小于非协调性惩罚,那么惩罚者会选择进行协调,此时 $\delta = 0$ 。因此,单个背叛者因遭受惩罚而付出的代价就表示为: $k = \begin{cases} p \sum_{i=1}^{ny} \alpha_i, & \delta = 0 \\ ny\delta(1-\delta)p \sum_{i=1}^{ny} \alpha_i, & 0 < \delta \leq 1 \end{cases}$ 。

理性的惩罚者总是通过选择 δ 而使惩罚作用最大化(如前所述,这更可能是演化的结果)。假设 k^* 为最大化的惩罚效果,那么 $k^* = \arg\max_{\delta} k$ 。这个数学问题的最优解可以这样求出:

(a) 当 $\delta > 0$ 时,惩罚者选择一致行动。此时,每位背叛者的损失为 $k = ny\delta(1-\delta)p \sum_{i=1}^{ny} \alpha_i$,理性的惩罚者选择 δ ,使得 $k^* = \arg\max_{\delta} k = \arg\max_{\delta} [ny\delta(1-\delta)p \sum_{i=1}^{ny} \alpha_i]$ 。假设最优的协调成本比例为 δ^* ,那么一阶条件可知 $ny(1-2\delta^*)p \sum_{i=1}^{ny} \alpha_i = 0$,即 $\delta^* = \frac{1}{2}$ 。此时 $k^* = \frac{1}{4}nyp \sum_{i=1}^{ny} \alpha_i$ 。

(b) 当 $\delta = 0$ 时,惩罚者投入协调的成本为 0,既然惩罚者之间不进行协调,惩罚就蜕变为独立惩罚的情景。此时,每位背叛者的损失为 $k^* = k = p \sum_{i=1}^{ny} \alpha_i$ 。

然后,理性的惩罚者会在这两种方式之间进行权衡,可以看出,当 $\frac{1}{4}nyp \sum_{i=1}^{ny} \alpha_i > p \sum_{i=1}^{ny} \alpha_i$ 即 $ny > 4$ 时,一致行动优于独立惩罚;而当 $ny < 4$ 时,独立惩罚优于一致行动;当 $ny = 4$ 时,二者无差异。因此, $k^* = \begin{cases} p \sum_{i=1}^{ny} \alpha_i, & ny \leq 4 \\ \frac{1}{4}nyp \sum_{i=1}^{ny} \alpha_i, & ny > 4 \end{cases}$ 。

同上一部分相同,单个合作者的期望收益仍然为 $P_c = b(x+y) - c$,单个惩罚者的收益仍然为 $P_p = b(x+y) - c - (1-x-y)np$,单个背叛者的收益仍然为 $P_b = b(x+y) - k$ 。将它们代入上文莫兰过程的演化框架,同样可求得 t 期合作者、背叛者和惩罚者的变化概率。我们在参数取值范围内设置初始参数,进行了数百次仿真,结果表现出很大程度上近似性。

图 4 显示了一个代表性仿真的结果,各参数的取值情况与非协调性惩罚情景下完全相同。这样,演化结果的差异在很大程度上可以归结为一致惩罚行动所致。从图 4 可以看出,在 100 万代仿真中,突变的惩罚者可以入侵这个群体,完成对背叛者的逆转,背叛者在人群中的平均比例与非协调性惩罚情形下相比明显降低,从 98.02% 下降为 6.25% (合作者的平均比例为 23.34%; 背叛者的平均比例为 6.25%; 惩罚者的平均比例为 70.40%)。可见,一致行动显著提高了群体的合作水平。

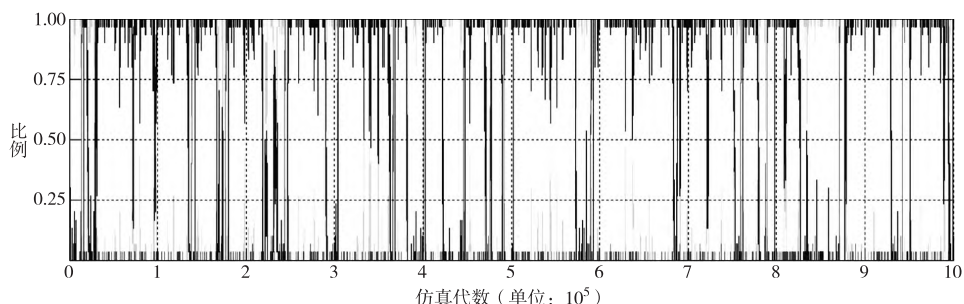


图 4 一致行动情形下的人群分布比例

为了更清楚地说明问题,我们专门截取了一些关键时期的仿真数据进行详细分析。在仿真运行初始阶段,与非协调性惩罚类似,大约 0—1500 代之间,该群体就可以完成合作者占据主导地位的局面,如图 5 所示;当合作者占据主导地位时,背叛者便可以入侵该群体,如图 6 所示。我们最为

关注的一个问题是惩罚如何涌现,即惩罚者如何从一个背叛者云集的情况下逆转,图 7 表示了该情形。可见,与非协调性惩罚不同,在适当的条件下,突变的惩罚者可以入侵这个群体,完成对背叛者的逆转。而当群体全部是惩罚者时,合作者又可以搭惩罚者的便车,再次入侵惩罚者占据主导地位,如图 8 所示。当合作者主导整个群体时,背叛者又卷土重来。这样,群体秩序表现出背叛者主导—惩罚者主导—合作者主导—背叛者主导这样的循环往复的情况,这种循环在 100 万次仿真中出现了十余次,如图 4 所示。

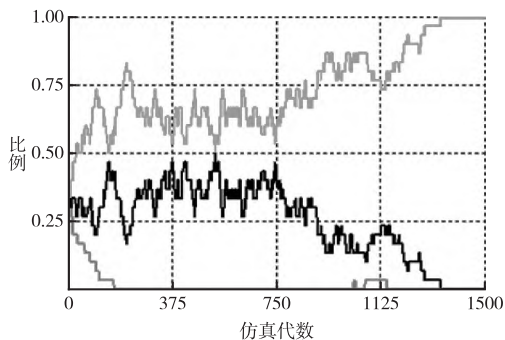


图 5 0—1500 代的仿真

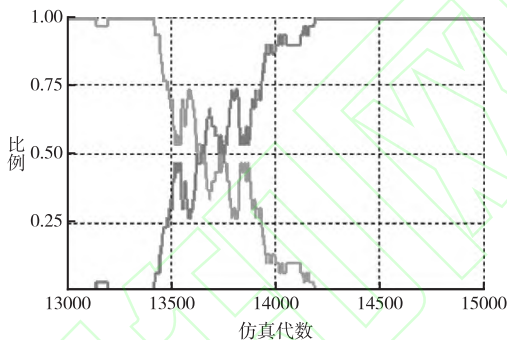


图 6 13000—15000 代的仿真

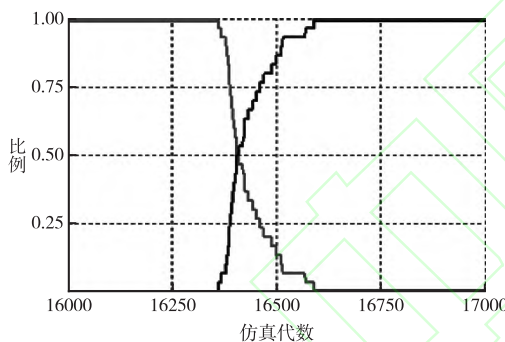


图 7 16001—17000 代的仿真

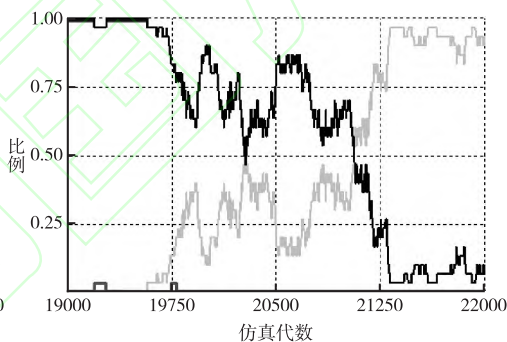


图 8 19001—424250 代的仿真

2. 代理惩罚

惩罚者的惩罚能力是不同的,如果能力较高的惩罚者能够代替能力较低的惩罚者进行惩罚的话,将是一个帕累托改善,因此,代理惩罚可以降低惩罚成本。一些人类学和历史学的证据(Sober & Wilson, 1997; Greif, 1993, 1994)表明,代理惩罚行为在人类社会十分普遍。比如,遇到背叛行为时,人们会首先考虑使用法律武器来惩罚背叛者和保护自己;老师、长者、领导等经常被要求充当惩罚者的角色,对他们所能影响的人群施加惩罚行为;某些人会向一些机构缴纳会费从而今后可以依靠该机构对背叛者进行惩罚。

代理惩罚又可细分为两种:一是雇佣惩罚,二是权威惩罚。二者虽然都存在着惩罚行为的委托,但前者仅指具有惩罚意愿的个体雇佣一个惩罚能力强的惩罚者进行惩罚,而后者则是指获得全体成员支持的权威实施惩罚。一般来讲,雇佣惩罚中群体成员之间是平等的,而权威惩罚中全体成员则存在等级关系;在雇佣惩罚中,委托方为获得超强的惩罚能力需要让渡一定的财富,而在权威惩罚中,委托方则是通过让渡财富或权力换取其对背叛者的有效惩罚行为;在雇佣惩罚中,代理方超强的惩罚能力来源于他们具备的某些特殊的生物性优势,比如较大的力量、强壮的身体等,而在权威惩罚中,代理方超强的惩罚能力更多地来自于其社会性优势,比如优越的社会和经济地位、某种特殊授权等;另外,雇佣惩罚一般存在着一种明确的雇佣行为,而在权威惩罚中,代理行为有时可

能并不明确,权威惩罚者会根据群体利益自主决定惩罚对象和惩罚强度。它们二者的具体区别如表 1 所示。

表 1 雇佣惩罚和权威惩罚的比较

区别内容	雇佣惩罚	权威惩罚
委托方	具有惩罚意愿的个体	全体成员
群体等级结构	成员之间平等	存在等级关系
代理方的报酬	物质利益	不仅限于物质利益,更多是权力的让渡
超强惩罚能力的来源	生物性优势	社会性优势
代理行为是否明确	一般存在明确的代理行为	有些情况明确,有些情况不明确

(1) 雇佣惩罚

为了方便对比,仍然沿用第二部分的分析框架,假定每个具有惩罚意愿的个体愿意对每一个背叛者付出惩罚成本 p 。一般来讲,代理方要想成功获得具有惩罚意愿的成员的委托,还必须通过一定的方式获得群体或社会的认同,其中的机制包括世袭、选举、建立声誉、武力竞争等。在本文中,我们为了研究方便,仅让声誉机制发挥作用。因此,与非协调性惩罚和一致行动不同,假定每次惩罚之前,每个惩罚者都会根据上一轮的博弈情况,选出一个惩罚能力最强的惩罚者进行委托,即每轮博弈后,群体成员可以观察到惩罚者的能力排名。然后,每个惩罚者把惩罚成本 p 交由该惩罚代理者进行惩罚,这样,每位背叛者付出的代价为 $k = nyp\alpha_{\max}$,其中 n_y 为惩罚者的人数, $\alpha_{\max}$ 为惩罚能力最强的惩罚者可以对背叛者造成的伤害倍数。那么,当存在雇佣惩罚时,单个合作者的期望收益为 $P_c = b(x+y) - c$, 单个惩罚者的收益为 $P_p = b(x+y) - c - (1-x-y)np$, 单个背叛者的收益为 $P_b = b(x+y) - k$ 。将它们代入上文莫兰过程的演化框架,同样可求得 t 期合作者、背叛者和惩罚者的变化概率。

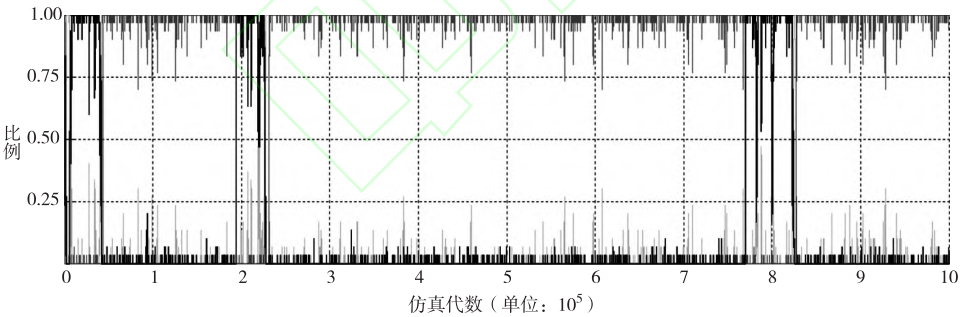


图 9 雇佣惩罚情形下的人群分布比例

图 9 显示了一个代表性仿真的结果,各参数的取值情况与非协调性惩罚情景下完全相同。可以看出,背叛者在人群中的平均比例与非协调性惩罚情形下相比明显降低,但不如一致性惩罚情景下低(合作者的平均比例为 2.20%; 背叛者的平均比例为 85.86%; 惩罚者的平均比例为 11.94%),可见,雇佣惩罚也提高了群体的合作水平。与一致惩罚情景下类似,在仿真运行初始阶段,大约 0—45000 代之间,该群体同样经历了从初始分布到被合作者主导再到被背叛者逆袭的局面,并且在适当的条件下,突变的惩罚者可以入侵这个群体,完成对背叛者的逆转。为了更清楚地说明问题,专门截取了 194001—232000 代的仿真情形,如图 10 所示,群体秩序表现出背叛者主导—惩罚者主导—合作者主导—背叛者主导这样的循环往复的情况。在大约 769001—829000 代之间,该群体又经历了一段相似的循环。

(2) 权威惩罚

在权威惩罚中,权威来自于全体成员的委托,因此,代理者作为权威可以向全体成员征“惩

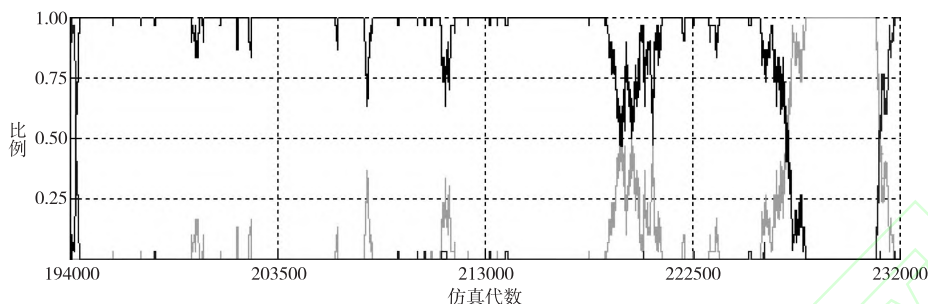


图 10 194001—232000 代的仿真

罚税”，即使背叛者也要缴纳，比如政府就可以向全体社会成员收缴税收。研究表明，群体具有趋向形成自我管理的机制，惩罚背叛者的权力一般被赋予特定的权威集中者，这可以提高惩罚的效力 (Henrich et al., 2010; Portes, 2010; Güth et al., 2007)。比如传统社会的村民会请求他们的首领来调节争端 (Gibson & Marks, 1995)、中世纪的欧洲商人创立了工会来维持商业秩序 (Greif, 1993)。King et al. (2009)、Rands et al. (2003)、Chase et al. (2002) 等人认为，角色分化和等级统治是大多数动物群体的组织原则，人类社会也不例外，集权化的制度可以更有效的克服协调问题和二次搭便车问题，从而比同辈惩罚更有效率。可见，人类社会的层级结构和权威在合作秩序维护中扮演了重要的角色。特别是在复杂的大规模社会，集权化惩罚和法定权威对维持社会的合作秩序更有意义，他们的权威来自所处社会阶层的位置 (Baldassarri & Grossman, 2011)。

我们仍然假定每个惩罚者对每一个背叛者付出的惩罚成本为 p ，并且为了对比，仍然假定权威的形成仅依靠声誉机制，这样每位背叛者付出的代价为 $k = np\alpha_{\max}$ 。此时，单个合作者与惩罚者的期望收益均为 $P_c = P_p = b(x + y) - c - (1 - x - y)np$ ，单个背叛者的收益为 $P_b = b(x + y) - (1 - x - y)np - k$ 。将它们代入上文莫兰过程的演化框架，同样可求得 t 期合作者、背叛者和惩罚者的变化概率。

图 11 显示了一个权威惩罚情景下的代表性仿真的结果，各参数的取值情况与非协调性惩罚情景下完全相同。可以看出，与非协调性惩罚和雇佣惩罚相比，背叛者在人群中的平均比例显著降低，惩罚者的比例显著提高（合作者的平均比例为 9.13%，背叛者的平均比例为 1.92%，惩罚者的平均比例为 88.95%）。

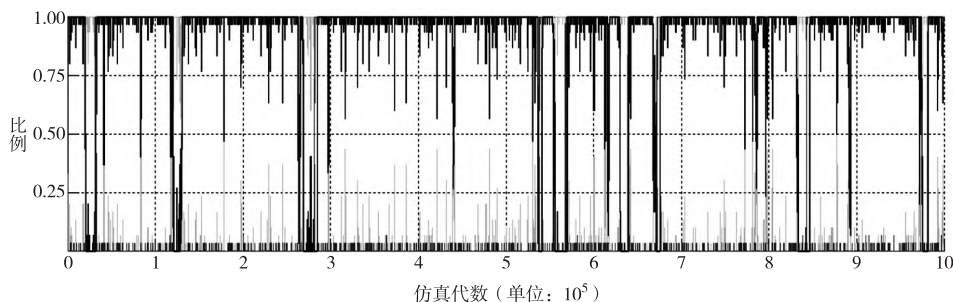


图 11 权威惩罚情形下人群分布比例

在仿真运行初始阶段，大约 0—200 代之间，该群体就可以完成惩罚者占据主导地位的局面。当惩罚者占据主导地位时，合作者便可以入侵该群体，如图 12 所示；而当合作者主导时，又会被背叛者逆袭，如图 13 所示；此时，突变的惩罚者又可以入侵这个群体，完成对背叛者的逆转，如图 14 所示。这样，群体秩序也会表现出背叛者主导—惩罚者主导—合作者主导—背叛者主导这样的循环往复的情况，这种循环在 100 万次仿真中出现了数次，如图 15 所示。

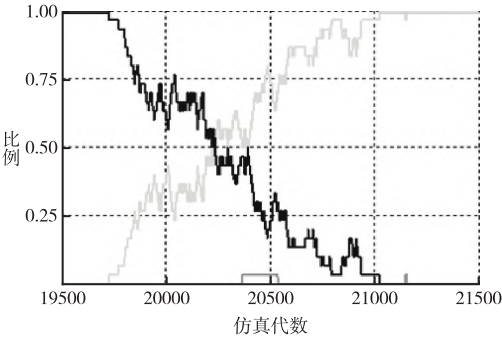


图 12 19501—21500 代的仿真

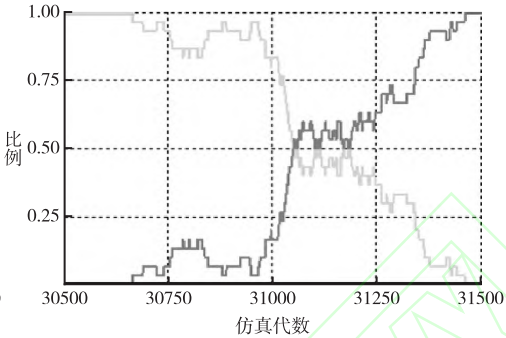


图 13 30501—31500 代的仿真

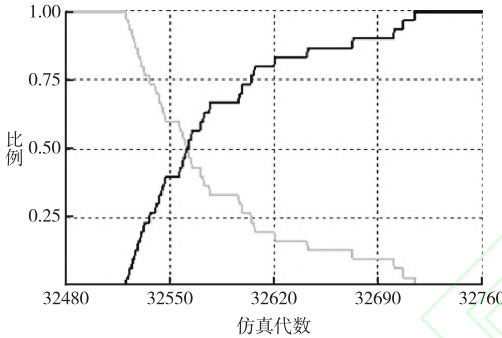


图 14 32501—32750 代的仿真

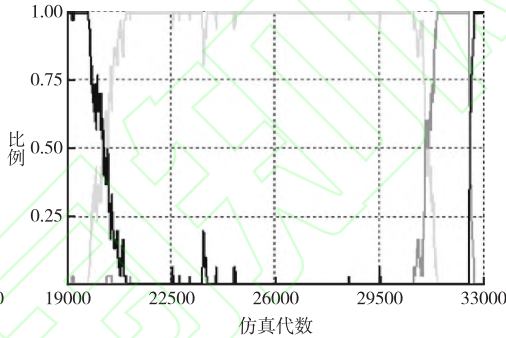


图 15 19000—33000 代的仿真

五、总 结

人类合作之谜堪称社会科学中最重要的难题之一。2005 年,《科学》杂志在创刊 125 周年之际,曾组织全世界最优秀的科学家提出了 25 个“驱动基础科学研究以及决定未来研究方向”的大难题,其中 23 个是纯粹的自然科学问题,比如“宇宙如何形成”、“意识的生物学基础是什么”等,只有两个问题同时还是社会科学问题,一个是“马尔萨斯的人口论观点会继续错下去吗”,另一个就是“人类合作行为如何演进”(黄少安和张苏,2013)。英国皇家学会主席 Robert May 也曾说过:“人类社会和其他动物种群中的合作行为如何演化和维持是演化生物学和更广泛的社会科学中最重要的未解之谜”(Colman,2006)。2016 年,《科学》杂志又对全世界最优秀的科学家进行了调查,“人类合作行为如何发展”再次入选 25 个最具挑战性科学难题之一。

虽然依靠惩罚可以维持群体合作的演化,但是惩罚会给惩罚者带来自然选择上的适应性劣势,从而带来新的演化之谜——降低惩罚者适应性的惩罚行为何以成功演化。我们认为,当面对背叛行为时,具有高级认知能力的人类既不会像其他物种那样转身离去,也不会简单、机械地对背叛者进行惩罚,而是会利用各种机制协调对付背叛者。在漫长的演化过程中,人类逐渐发展出了各种巧妙的协调性惩罚行为,这些协调无疑会降低惩罚成本。当它们改变惩罚行为的成本劣势,使其达到成功演化的阈值范围时,反对惩罚的选择性压力就消失,从而可以支持利他性惩罚在大规模群体中演化。本文不再假设惩罚者是同质的,而是假设其惩罚能力服从正态分布,这样便可以讨论各种协调性惩罚对群体合作秩序演化的影响。在此基础上构建动态演化模型,并根据文献资料确定了合适的参数取值范围,模拟了人类祖先的一个典型狩猎采集社会,对非协调性惩罚、一致行动以及代理惩罚(包括雇佣惩罚和权威惩罚两种情况)等情景下的群体合作的演化进行了仿真分析。在这几种情景下,设置各参数的取值完全相同,这样,演化结果的差异在很大程度上可以归结为不同的

协调机制所致。仿真结果表明,非协调性惩罚很难维持较大规模群体中合作的演化,而协调性惩罚会提高群体合作水平,特别是在一致行动和权威惩罚的情景中,群体在大部分时间内都可以实现较高程度的合作。

必须要说明的是,虽然许多文献都对降低自身适应性的惩罚行为何以成功演化这一“二阶搭便车”问题进行了研究,但解释依然还不完美。本文只是提出了一个思路,有创新也有缺陷。协调虽然可以降低惩罚成本,但是,协调本身也需要成本。因此,惩罚者会在协调和不惩罚之间进行权衡,这需要进行后续更加深入的研究。另外,我们从来不认为非协调性惩罚状态在历史上真实存在,而只是将其作为一个参照系,来讨论协调性惩罚的作用,就像物理研究的真空状态。

参考文献

- 董志强、李伟成,2019:《禀赋效应和自然产权的演化:一个主体基模型》,《经济研究》第1期。
- 弗朗西斯·福山,2015:《大断裂:人类本性与社会秩序的重建》,广西师范大学出版社。
- 黄少安、张苏,2013:《人类的合作及其演进:研究综述和评论》,《中国社会科学》第7期。
- 韦倩,2009:《增强惩罚能力的若干社会机制与群体合作秩序的维持》,《经济研究》第10期。
- 周业安、宋紫峰,2008:《公共品的自愿供给机制:一项实验研究》,《经济研究》第7期。
- Axelrod, R., and W. D. Hamilton, 1981, “The Evolution of Cooperation”, *Science*, 211, 1390—1396.
- Baldassarri, D., and G. Grossman, 2011, “Centralized Sanctioning and Legitimate Authority Promote Cooperation in Humans”, *Proceedings of the National Academy of Sciences of the United States of America*, 108, 11023—11027.
- Boehm, C. H., 1993, “Egalitarian Behavior and Reverse Dominance Hierarchy”, *Current Anthropology*, 34, 227—254.
- Boehm, C. H., 2007, *Conscience Origins, Sanctioning Selection, and the Evolution of Altruism in Homo Sapiens*, Department of Anthropology, University of Southern California.
- Bowles, S., and H. Gintis, 2004, “The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations”, *Theoretical Population Biology*, 65, 17—28.
- Bowles, S., and H. Gintis, 2011, *A Cooperative Species: Human Reciprocity and Its Evolution*, Princeton, New Jersey Oxford: Princeton University Press.
- Boyd, R., and P. J. Richerson, 1985, *Culture and the Evolutionary Process*, Chicago: University of Chicago Press.
- Boyd, R., and P. J. Richerson, 1992, “Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups”, *Ethology and Sociobiology*, 13, 171—195.
- Boyd, R., H. Gintis, and S. Bowles, 2010, “Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare”, *Science*, 328, 617—620.
- Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson, 2003, “The Evolution of Altruistic Punishment”, *Proceedings of the National Academy of Sciences of the United States of America*, 100(6), 3531—3535.
- Chase, I., C. Tovey, D. Spangle, M. Manfredonia, 2002, “Individual Differences versus Social Dynamics in the Formation of Animal Dominance Hierarchies”, *Proceedings of the National Academy of Sciences of the United States of America*, 99, 5744—5749.
- Colman, A. M., 2006, “The Puzzle of Cooperation”, *Nature*, 440, 744—745.
- Darlington, P. J., 1975, “Group Selection, Altruism, Reinforcement and Throwing in Human Evolution”, *Proceedings of the National Academy of Sciences of the United States of America*, 72, 3748—3752.
- Dawkins, R., 1986, *The Blind Watchmaker*, New York: W. W. Norton.
- De Waal, F., 1989, *Chimpanzee Politics: Power and Sex Among Apes*, Baltimore: Johns Hopkins University Press.
- Fifer, F. C., 1987, “The Adoption of Bipedalism by the Hominids: A New Hypothesis”, *Human Evolution*, 2(2), 135—147.
- Fowler, J. H., 2005, “Altruistic Punishment and the Origin of Cooperation”, *Proceedings of the National Academy of Sciences*, 102, 7047—7049.
- Gächter, S., E. Renner, and M. Sefton, 2008, “The Long-run Benefits of Punishment”, *Science*, 322(5907), 1510—1512.
- Gibson, C. C., and S. A. Marks, 1995, “Transforming Rural Hunters into Conservationists: An Assessment of Community-Based Wildlife Management Programs in Africa”, *World Development*, 23, 941—957.
- Gintis, H., 2000, “Strong Reciprocity and Human Sociality”, *Journal of Theoretical Biology*, 206, 169—179.

- Gintis, H. , E. A. Smith, and S. Bowles, 2001, “Costly Signaling and Cooperation”, *Journal of Theoretical Biology*, 213, 103—119.
- Greif, A. , 1993, “Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders’ Coalition”, *American Economic Review*, 83, 525—548.
- Greif, A. , 1994, “Cultural Beliefs and The Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies”, *Journal of Political Economy*, 102, 912—950.
- Güth, W. , M. V. Levati, M. Sutter, and E. van der Heijden, 2007, “Leading by Example with and without Exclusion Power in Voluntary Contribution Experiments”, *Journal of Public Economics*, 91, 1023—1042.
- Harari, Y. N. , 2014, *Sapiens: A Brief History of Humankind*, London: Harvill Secker.
- Hauert, C. , A. Traulsen, H. Brandt, M. A. Nowak, and K. Sigmund, 2007, “Via Freedom to Coercion: The Emergence of Costly Punishment”, *Science*, 316, 1905—1907.
- Henrich, J. , and R. Boyd, 2001, “Why People Punish Defectors: Weak Conformist Transmission Can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas”, *Journal of Theoretical Biology*, 208, 79—89.
- Henrich, J. , J. Ensminger, R. McElreath, A. Barr, C. Barrett, A. Bolyanatz, J. C. Cardenas, M. Gurven, E. Gwako, N. Henrich, C. Lesorogol, F. Marlowe, D. Tracer, and J. Ziker, 2010, “Markets, Religion, Community Size, and the Evolution of Fairness and Punishment”, *Science*, 327(5972), 1480—1484.
- Helbing, D. , A. Szolnoki, M. Perc, and G. Szabó, 2010, “Evolutionary Establishment of Moral and Double Moral Standards through Spatial Interactions”, *PLoS Computational Biology*, 6(4), e1000758.
- Jaffe, K. , and L. Zaballa, 2010, “Co-operative Punishment Cements Social Cohesion”, *Journal of Artificial Societies and Social Simulation*, 13(3), 4.
- King, A. J. , Dominic D. P. Johnson, and M. Van Vugt, 2009, “The Origins and Evolution of Leadership”, *Current Biology*, 19, 911—916.
- Marlowe, F. W. , 2005, “Hunter-gatherers and Human Evolution”, *Evolutionary Anthropology*, 14, 54—67.
- Nowak, M. A. , 2006, *Evolutionary Dynamics: Exploring the Equations of Life*, Cambridge: Harvard University Press.
- Nowak, M. A. , and K. Sigmund, 1998, “Evolution of Indirect Reciprocity by Image Scoring”, *Nature*, 393, 573—577.
- Nowak, M. A. , and R. M. May, 1992, “Evolutionary Games and Spatial Chaos”, *Nature*, 359(6398), 826—829.
- Panchanathan, K. , and R. Boyd, 2004, “Indirect Reciprocity Can Stabilize Cooperation without the Second-Order Free-Rider Problem”, *Nature*, 432, 499—502.
- Portes, A. , 2010, *Economic Sociology: A Systematic Inquiry*, Princeton University Press; Princeton, NJ.
- Pugh, G. E. , 1977, *The Biological Origin of Human Values*, New York: Basic Books.
- Rands, S. A. , G. Cowlshaw, R. A. Pettifor, J. M. Rowcliffe, and R. A. Johnstone, 2003, “Spontaneous Emergence of Leaders and Followers in Foraging Pairs”, *Nature*, 423(6938), 432—434.
- Sasaki, T. , and S. Uchida, 2013, “The Evolution of Cooperation by Social Exclusion”, *Proceedings of the Royal Society B*, 280(1752), 1—7.
- Sethi, R. , and E. Somanathan, 1996, “The Evolution of Social Norms in Common Property Eesource Use”, *American Economic Review*, 86(4), 766—788.
- Sober, E. , and D. S. Wilson, 1997, *Unto Others—The Evolution and Psychology of Unselfish Behavior*, Cambridge, MA: Harvard University Press.
- Szolnoki, A. , and M. Perc, 2013, “Evolving Cooperation”, *Journal of Theoretical Biology*, 299, 1—8.
- Traulsen, A. , N. Shores, and M. Nowak, 2008, “Analytical Results for Individual and Group Selection of Any Intensity”, *Bulletin of Mathematical Biology*, 70, 1410—1424.
- Trivers, R. L. , 1971, “The Evolution of Reciprocal Altruism”, *Quarterly Review of Biology*, 46, 35—57.
- Wade, N. , 2009, *The Faith Instinct: How Religion Evolved and Why It Endures*, 18—37.
- Wedekind, C. , and M. Milinski, 2000, “Cooperation through Image Scoring in Humans”, *Science*, 289, 850—852.
- Wiessner, P. , 2005, “Norm Enforcement among the Ju/’hoansi Bushmen”, *Human Nature*, 16(2), 115—145.
- Wrangham, R. , and D. Peterson, 1996, *Demonic Males: Apes and the Origins of Human Violent*, Boston: Houghton Mifflin.

Coordinated Punishment and the Evolution of Human Cooperation

WEI Qian^a, SUN Ruiqi^a, JIANG Shuguang^b and YE Hang^b

(a: Center for Economic Research, Shandong University; b: School of Economics,

Center for Economic Behavior Decision-making (CEBD), Zhejiang University of Finance and Economics)

Summary: Although punishment can promote and sustain the evolution of cooperation, it leads to adaptive natural selection disadvantages for the punishers; thus the evolution of punishment is a mystery. The literature explains this problem mainly from the perspective of compensation, that is, the benefits from punishment behavior can offset its adaptive disadvantages to the punishers. While this stream of literature provides valuable insights into the mystery of human cooperation, its shortcomings are obvious. First, it assumes that the cost of punishment is not very high. Frustratingly, however, this assumption is not always well established in reality. The punishers may be not competent enough to defeat the betrayers in a fight, and betrayers often take revenge, which makes the cost of punishment extremely high. It is obvious that when considering the high cost of punishment in certain cases, any explanation based on a compensation mechanism is not convincing. Second, it is usually insufficient to describe and analyze the punishment behavior itself. We believe that human beings with advanced cognitive abilities do not turn away like other species when facing betrayal, and do not merely punish the betrayers mechanically. Instead they use various coordinated mechanisms to confront the betrayers. In this sense, a specialized and in-depth analysis of punishment behavior helps to provide a clearer understanding of the true evolution of human cooperation.

Unlike the literature based on the compensation mechanism, the present study focuses on the punishment behavior itself, which is a more realistic simulation of evolutionary history. We propose that during the long evolutionary process, human beings gradually developed various coordinated punishment behaviors, which help to reduce the cost or increase the severity of punishment. When this advantage surpasses the cost disadvantage of punishment to reach the threshold condition for the successful evolution of punishment, the selective pressure against punishment disappears, which can sustain the evolution of altruistic punishment in large-scale groups. We construct a dynamic evolutionary model and simulate a typical hunter-gatherer society of human ancestors to investigate the impact of coordinated punishment on the evolution of cooperation. We also compare and analyze three different mechanisms of coordinated punishment and their impact on the evolution of cooperation. The simulation results show that it is difficult for uncoordinated punishment to sustain the evolution of cooperation in large-scale groups, whereas coordinated punishment promotes group cooperation, especially in the case of collective and authoritative punishment, where the group can maintain a high level of cooperation most of the time.

The present study makes important contributions to the related literature. First, we contribute to the large volume of literature highlighting the role of punishment in the successful evolution of cooperation. We propose a new perspective on coordinated punishment, which has two notable merits. First, it is more realistic, and second, coordinated punishment can sustain the stable evolution of cooperation even when a single punisher faces a very high cost of punishment. Second, we further subdivide coordinated punishment into two types: collective and principal-agent punishment. Principal-agent punishment can be further divided into employment punishment and authoritative punishment. We then test and compare the impact of the three types of coordinated punishment on the evolution of cooperation. Third, we differ from the previous literature, such as Jaffe & Zaballa (2010) and Boyd et al. (2010), in our model settings where punishers are homogeneous agents, while allowing their punishment competence to be heterogeneous and follow a normal distribution. This assumption is more realistic and makes it possible to discuss and compare the effects of the three types of coordinated punishment on group cooperation.

Keywords: Cooperation; Coordinated Punishment; Evolution; Simulation

JEL Classification: D01, D02, D70, D74

(责任编辑:陈 辉)(校对:南 山)