

# 合作的逻辑：基于强欺骗策略的演化分析<sup>①</sup>

黄 璜

(北京大学政府管理学院, 北京 100871)

**摘要:** 如何解释非亲缘关系之间的合作现象? 合作演化( evolution of cooperation) 研究在多人囚徒困境博弈模型基础上提出一系列理论上的解释. 作者利用“基于主体建模”方法建立了基于惩罚策略的合作演化模型. 模型中有两种惩罚策略. 一种是强合作策略, 即合作且惩罚不合作者; 另一种是强欺骗策略, 即不合作且惩罚不合作者. 模拟结果说明, 虽然强合作策略在个体层面有利合作, 却可能在宏观上破坏合作秩序; 强欺骗策略在微观上不利于合作, 但在整体上, 即使是在较大规模的社会中, 可能有助于社会合作的形成, 因为该策略主体在形成不稳定联盟或多个联盟时会导致出现合作与非合作并存的混合策略均衡, 也即虽然不能在最优水平, 却在次优水平上实现稳定的社会合作.

**关键词:** 合作演化; 惩罚; 强欺骗策略; 不稳定联盟; 大规模社会合作

**中图分类号:** N945. 12; O225 **文献标识码:** A **文章编号:** 1007 - 9807( 2013) 09 - 0001 - 08

## 0 引 言

“囚徒困境”博弈以极其简单的逻辑证明了个人理性与集体理性发生冲突的可能性. 然而现实社会中却能观察到大量合作现象. 早期“亲缘选择( kin selection) ”理论认为, 合作是基因延续的需要<sup>[1]</sup>. “互惠”理论认为, 在非亲缘条件下合作是为了保证在长期交往中能够实现持续的互惠<sup>[2]</sup>. Axelrod<sup>[3]</sup>的计算机模拟实验证明, “一报还一报( tit for tat) ”, 即“互惠”, 是“演化稳定策略( evolution stable strategy) ”<sup>[4]</sup>. 直接“互惠”无法解释一次性合作, “间接互惠”理论则提出, 陌生人之间的合作是一条合作链, 你帮助我, 是因为他曾帮助你, 或因为我会帮他; 人们为维护“声誉”、“地位”而与陌生人合作<sup>[5]</sup>. 也有人指出, 合作是一种用于表明自己质量和动机的高成本信号<sup>[6]</sup>. Nowak<sup>[7]</sup>等的模拟实验证明, 当个体数量较少时可以通过“间接互惠”实现合作均衡. 其后一系列

文献对该模型进行了修正<sup>[8-12]</sup>.

互惠模型在较大规模社会中的作用有限, 有效地“惩罚”更具解释力, 因为惩罚能降低不合作者的净福利, 从而使“不合作”成为得不偿失的选择<sup>[13]</sup>. “强互惠( strong reciprocity) ”理论提出“强互惠者”不仅会与其他人合作, 而且愿意支付本来惩罚不合作者<sup>[14]</sup>. 计算机模拟实验揭示在“公共产品博弈”( 多人重复“囚徒困境”博弈) 实验中, 强互惠者在一定条件下有利于合作的演化<sup>[15]</sup>. 不过“强互惠”策略不是演化稳定策略, 因为它同时支付合作和惩罚的成本. 强互惠理论将“强互惠”动机归结于“利他性”, 试图证明这种“利他性”的生理存在<sup>[16]</sup>. 另一些模型则将动机归因于对自身得失的反应<sup>[17-19]</sup>. 在“强互惠”理论基础上提出了“文化与基因共演化( culture and gene co-evolution) ”理论<sup>[20-21]</sup>、“双继承( dual inheritance) ”理论<sup>[22]</sup>等, 从社群差异与竞争的角度解释“大规模社会”中的合作演化<sup>[23-24]</sup>.

① 收稿日期: 2011 - 06 - 20; 修订日期: 2013 - 03 - 03.

基金项目: 国家社会科学基金资助项目( 09CZZ009) .

作者简介: 黄 璜( 1978—), 男, 江苏南通人, 博士, 副教授. Email: huanghuang@ pku. edu. cn

惩罚本身也面临合作困境,那些没有承担惩罚成本的人搭了便车.这便涉及新一轮“惩罚”与“合作”,这样便形成循环:在每次“循环”中,个体都需要选择新策略,而每次选择意味着又有新策略的出现.虽然研究证明,这种多次循环有利于维持合作稳定<sup>[25]</sup>,不过实证研究并不支持<sup>[26]</sup>.此外“奖赏”在合作中的作用曾被认为逊色于“惩罚”<sup>[27-32]</sup>,但近来也有不同观点<sup>[33]</sup>.“奖赏”实质上是一种“互惠”,有鉴别的奖赏即为惩罚.有实验证明,为保证互惠,人们会增强在公共产品合作中的投入以保持声誉<sup>[34,35]</sup>.

合作演化研究已在国内学界引起关注<sup>[36-40]</sup>.若干模型证明个体间关系不对称性或合作者网络有助于维持合作关系的稳定<sup>[41,42]</sup>.基于演化博弈建立的收益和惩罚共享博弈模型对群体工作中合作行为进行多智能体模拟研究<sup>[43]</sup>.此外有学者提出职业化强互惠概念,但没有给出数理证明<sup>[44]</sup>.

合作演化研究的基本方法是以囚徒困境博弈为基础建立形式化模型并实现模拟.合作演化模型是对人类社会的高度抽象,认为社会是能够自组织、自适应、非线性的“复杂系统”.软件技术在20世纪50年代进入社会科学,其中“基于主体建模(agent-based modeling)”在为社会科学的“非线性”的复杂问题建模有突出作用<sup>[45,46]</sup>,尤其对研究合作演化博弈有很好的适用性.

本文利用“基于主体建模”方法,构建基于“惩罚”的合作演化模型.一个典型的合作博弈模型涉及合作和欺骗两种策略.而在“惩罚”这个层面上,由表1可知,合作者和欺骗者又有两种选择:惩罚或不惩罚.因此可以得到4种策略,即强合作、弱合作、弱欺骗和强欺骗.过去的文献主要关注前3种策略.实际上,欺骗者越多,它们能“剥削”的收益就越少;如果通过支付一定成本让其他欺骗者参加合作,便可获得更多收益,此即强欺骗策略.有文献证明,在特定条件下4种策略都可能成为个人的最佳选择<sup>[47]</sup>.本文试图证明,强欺骗策略在一定条件下有助于形成合作,尤其是大规模社会合作.

表 1 基于“惩罚”的策略选择矩阵

Table. 1 The strategy selection matrix based of punishment

	合作	欺骗
惩罚	强合作	强欺骗
不惩罚	弱合作	弱欺骗

### 1 合作演化的模型

假设存在完全封闭的“社会-空间系统”,包含若干个体和环境.个体共同生产某种公共资源,收益在所有个体间均分.个体可参加(合作)或不参加(不合作)生产.假设个体数量为 $N(N \geq 2)$ ,其中合作者为 $n$ .生产中个人投入成本 $c$ (常数),收益为 $F(n)$ .公共资源供给模型为<sup>[47]</sup>

$$L = 1 - (1 - PC)^E \tag{1}$$

其中 $PC = n/N$ ;  $L$ 为公共资源的开发水平,值域为 $[0, 1]$ (0:无公共资源,1:最高水平). $PC = 0$ ,则 $L = 0$ , $PC = 1$ ,则 $L = 1$ ;当 $0 < PC < 1$ 时 $PC$ 和 $L$ 的关系取决于 $E$ .若 $E = 1$ , $L = PC$ ,边际效益不变.若 $E \neq 1$ ,生产函数是非线性函数:当 $E > 1$ 时,边际收益递减,当 $E < 1$ 时,边际收益递增,3种函数有不同性质<sup>[48]</sup>.

本模型假设存在最优供应水平 $M$ ,每个个体收益为

$$F(n) = LM = \left[ 1 - \left( 1 - \frac{n}{N} \right)^E \right] M \tag{2}$$

为了简化讨论,仅考虑较为普遍的情况,即边际效益递减.令 $E = 2$ ,合作者净收益为

$$Y_{co} = LM - c = \frac{nM(2N - n)}{N^2} - c \tag{3}$$

欺骗者净收益为

$$Y_{de} = \frac{nM(2N - n)}{N^2} \tag{4}$$

环境是一个 $X \times X$ 的矩阵空间<sup>②</sup>.在每个时刻,每个个体唯一地占用空间中一个单元,并采用“摩尔邻居”模式,即每个单元有上下左右和对角8个邻居.每个单元中随机地以速率 $r$ 出产资源.个体为寻找更多资源会向四周移动.假设公共资源和随机资源同质,可累加.

假设个体拥有相同“寿命”(L),“死亡”时

② 为规避边界效应,空间上、下、左、右边界相互衔接,实际上是个“环形曲面(torus)”.

将繁衍后代. 后代数量取决于个体资源存量“适存度(fitness)” $f^{\text{③}}$ .  $f$ 数值排前1/3的个体有2个后代, 后1/3没有后代, 否则为1个. 诞生时初始“适存度”为 $if$ , 父辈将且仅将其策略“遗传”给子女, 但在小概率 $\varphi$ 范围内随机变异成另一种策略. 假设个体“新陈代谢”消耗资源为 $m$ , 视野范围为 $v$ , 移动速度为 $s$ . 由于个体是运动的, 所以观察范围是动态的. 任何主体都无需了解策略完整分布, 也无需进行复杂计算. 当欺骗者出现在惩罚者视野范围内, 后者会按“发现概率” $\theta(\theta < 1)$ 惩罚. 一旦惩罚, 欺骗者 $f$ 减少 $\rho$ , 惩罚者 $f$ 减少 $\sigma$ ,  $\rho > \sigma$ .

强欺骗者两两遭遇时会相互惩罚. 考虑4种情景. 第1, “无联盟”, 即任由互相惩罚. 第2, “单联盟”, 强欺骗者之间放弃惩罚, 形成“联盟”. 第3, “有条件联盟”, 联盟是不稳定的, 假设因某种原因不一定达成妥协, 存在妥协成功概率 $\eta$ . 第4, “双联盟”, 假设存在某种标志(如生理、族群、阶级、意识形态等)将强欺骗者分为两个联盟, 联盟内能妥协, 但相互间无法妥协. 这4种情景相互关联, “无联盟”可看作 $\eta = 0$ 或联盟容量为1, “单联盟”则是 $\eta = 1$ 或联盟个数为1.

## 2 模拟实验及其结果

利用 REPAST 开发平台开发了一个软件系统. 系统通过“时钟”同步主体行动, “时钟”是离散的, 每个时间段称作“时间单元”. 在每个时间单元, 主体为公共生产提供1次劳动并分享收益, 寻找1次随机资源, 消耗旧资源, 惩罚者惩罚视野范围内的欺骗者. 主体活动是同时进行的, 不存在任何外生力量安排其活动顺序. 但由于单处理器无法并行计算, 因此在每个时间单元初始, 主体被随机排序, 这样不会因行动顺序而导致显著差异.

实验结果有两个评价指标“稳定性(stability)”和“演化性( evolvability)”. 假设一个群体中所有个体都采取某策略, 而现有一小群采纳另一策略的个体进入该群体, 那么如果原策略

在收益上比新策略好, 便可阻止新策略“入侵”, 称原策略具有演化“稳定性”(该策略是“演化稳定策略”); 如果新策略能成功“入侵”, 称其具有“演化性”<sup>[49]</sup>. 合作演化研究关注合作策略是否具有“稳定性”和“演化性”. “稳定性”指少数个体的“搭便车”不会因为被模仿而成为普遍行为; “演化性”指在普遍互不信任, 无法合作时, 少数合作者能否推动实现普遍的合作关系.

实验参数如表2所示. 各情景模拟10次, 每次按需分别运行1万、2.5万和5万个时间单元, 每500个时间单元计算1次主体数量, 结果取10次实验平均值.

表2 实验的基本参数

Table. 2 Experiment parameters

$X$	40	$N$	180	$E$	2
$Ls$	50	$\phi$	1%	$M$	0.5
$if$	20	$r$	1%	$c$	0.1
$m$	在[0.1 0.3] 正态分布, 小数取1位				
$v$	在[2 4] 均匀分布, 取整数				
$s$	在[1 3] 均匀分布, 取整数				
$\rho$	1	$\sigma$	0.1		

首先, 假设只存在弱合作者、弱欺骗者和强欺骗者. 初始时主体数量平均分布. 无联盟和单联盟模拟结果如图1. 在无联盟情景中, 弱欺骗者在演化伊始就表现出显著的演化稳定性. 强欺骗者不具备演化优势, 对结果几乎没有影响. 在单联盟情景中, 强欺骗策略则能够稳定演化, 这时它和弱欺骗策略类似. 其他实验证明, 若弱欺骗者在初始时占绝对优势, 强欺骗策略也不能演化成功. 总之, 无联盟和单联盟的演化结果都是非合作均衡.

其次, 有条件联盟情景结果如图2. 如果妥协成功概率 $\eta = 80\%$ , 弱合作者数量迅速提高到90%左右, 并保持到实验结束; 强欺骗者在10%左右稳定下来; 弱欺骗者则被迅速“挤出”. 当 $\eta = 95\%$  则弱合作者数量保持在55%左右, 强欺骗者占45%左右. 这样形成一种弱合作者和强欺骗者并存的混合策略均衡.

③ 适存度是进化论词汇, 描述生物个体或基因的生存机会: 适存度越大, 个体自然适应力越强, 就可能有更多后代.

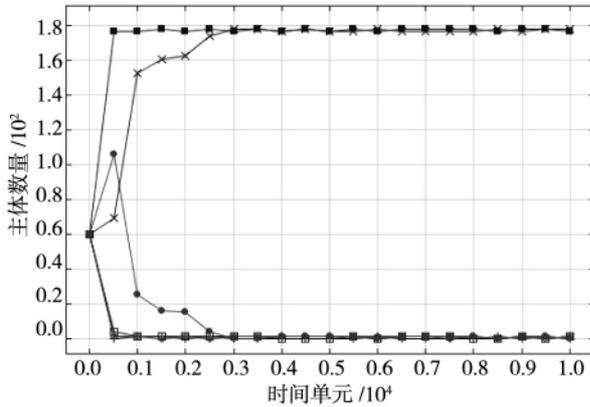


图1 无联盟和单联盟情景实验, 初始时策略平均分布

无联盟: ●弱合作者, ×弱欺骗者, □强欺骗者;  
单联盟: ◇弱合作者, +弱欺骗者, ■强欺骗者.

Fig. 1 Non-union and single union expt., equal types initially.

Non-union: ●weak C, ×weak D, □strong D;  
Single union: ◇weak C, +weak D, ■strong D.

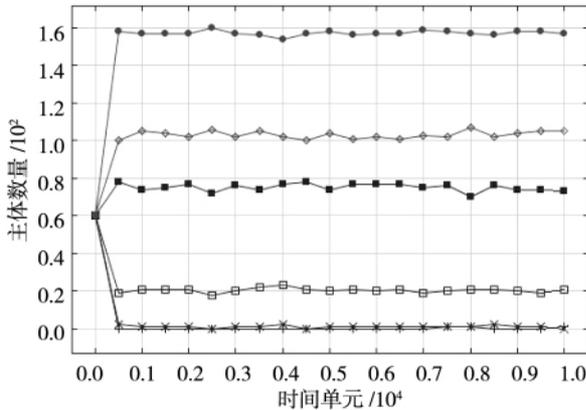


图2 有条件联盟情景实验, 初始类型平均分布

$\eta = 80\%$ : ●弱合作者, ×弱欺骗者, □强欺骗者;  
 $\eta = 95\%$ : ◇弱合作者, +弱欺骗者, ■强欺骗者.

Fig. 2 Conditional union expt., equal types initially.

$\eta = 80\%$ : ●weak C, ×weak D, □strong D;  
 $\eta = 95\%$ : ◇weak C, +weak D, ■strong D.

再次, 双联盟情景结果如图3. 如果令发现概率  $\theta = 50\%$ , 弱合作者数量很快达到最大值, 之后逐步下降, 弱欺骗者数量逐渐上升. 与无联盟情景相比, 弱欺骗者演化速度大大降低, 但最终仍进入弱欺骗策略均衡<sup>④</sup>. 当  $\theta = 20\%$  时, 又出现弱合作和强欺骗并存的混合策略均衡, 后者占15%左右. 当  $\theta = 5\%$ , 同样是混合策略均衡, 强欺骗者比例接近50%. 混合策略均衡说明, 虽然合作不是

所有主体的选择, 但是其数量可维持在稳定水平上. 混合策略是种演化稳定策略. 现将系统扩大到  $X \times X = 100 \times 100$ ,  $N = 1200$ ; 设  $\theta = 50\%$ ,  $\eta = 80\%$  (有条件联盟) 或者  $\theta = 20\%$  (双联盟). 结果如图4, 系统也进入混合策略均衡, 弱合作者数量占到85%以上.

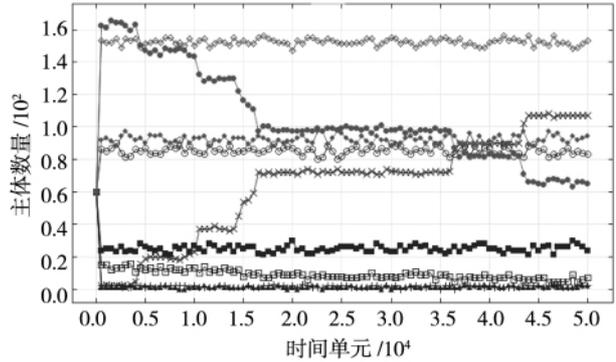


图3 双联盟实验, 初始策略平均分布

$\theta = 50\%$ : ●弱合作者, ×弱欺骗者, □强欺骗者;  
 $\theta = 20\%$ : ◇弱合作者, +弱欺骗者, ■强欺骗者;  
 $\theta = 5\%$ : ◆弱合作者, ▲弱欺骗者, ○强欺骗者.

Fig. 3 Duo-union expt., equal types initially.

$\theta = 50\%$ : ●weak C, ×weak D, □strong D;  
 $\theta = 20\%$ : ◇weak C, +weak D, ■strong D;  
 $\theta = 5\%$ : ◆weak C, ▲weak D, ○strong D.

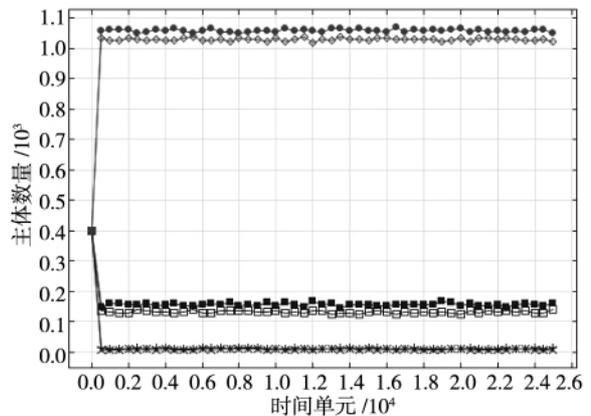


图4 大规模条件下有条件联盟和双联盟实验

有条件联盟: ●弱合作者, ×弱欺骗者, □强欺骗者;  
双联盟: ◇弱合作者, +弱欺骗者, ■强欺骗者

Fig. 4 Conditional union and duo-union in large-scale expt.

Condi. union: ●weak C, ×weak D, □strong D;  
Duo-union: ◇weak C, +weak defector, ■strong D.

④ 图中显示3种策略并存是因为存在两种实验结果, 即弱欺骗策略均衡、弱合作和强欺骗并存的混合策略均衡. 虽然在规定实验时间内可能保持混合策略状态, 但这种状态不稳定.

最后,引入强合作者,如图5,若无强欺骗者,强合作者无助于实现普遍合作.而如图6(有条件联盟)和图7(双联盟),这时合作演化能力弱于强欺骗者单独存在时的情况,只有 $\eta$ 很高(有条件联盟)或 $\theta$ 很低(双联盟)时才出现混合策略均衡,否则进入弱欺骗均衡.

### 3 对模拟结果的讨论

以上结果说明,当强欺骗者处于完全分散(无联盟)或完全联合(单联盟)两个极端状态时,系统最终形成非合作均衡;而当处于部分联合状态时,可以形成弱合作策略与强欺骗策略并存的混合策略均衡.直接的解释是,第1,在单联盟中,强欺骗者无偿占用合作收益却免受惩罚,而弱欺骗者要被惩罚,因此最佳策略是加入强欺骗者的“联盟”.第2,当联盟中出现“分歧”时,强欺骗者之间的冲突导致其收益下降;但只要妥协可能性仍较高,或惩罚意愿较弱,强欺骗者就仍有条件控制弱欺骗者;由于欺骗者收益都会下降,弱合作者即使被剥削,也能使相对稳定的收益流“涌现”出来,这样形成混合策略均衡.第3,如果分歧很大,过于频繁的“相互攻伐”导致强欺骗者收益迅速下降,生存受到威胁,控制力大大降低,这时弱欺骗者优势渐显;如果妥协完全无法成功或联盟被原子化,强欺骗者的作用便完全消失了.而强合作者由于自身收益结构的缺陷,没有演化稳定性.当4种策略并存,强合作者对强欺骗者的惩罚增加了后者的负担,因此只在妥协概率很高或惩罚意愿很弱时,才能实现混合策略均衡,此时强合作者的作用可忽略,等同于没有强合作者时的情况.

总结以下两点.第1,强欺骗者虽仍是搭便车者,但在宏观上,即使在较大社会范围内可能发挥“合作稳定器”的作用.相反,强合作者虽从个体上看有利于合作,但却可能破坏合作秩序.第2,社会系统有可能自发演化出合作调节机制,即合作与非合作并存的策略均衡,这样使得虽不能在最优水平,但可在次优水平上实现合作收益的最大化.

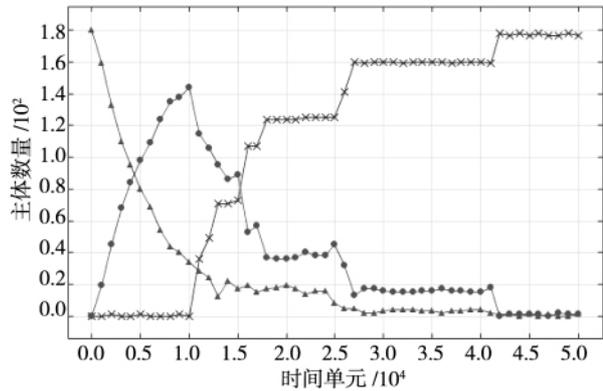


图5 强合作策略演化, ●弱合作者, ▲强合作者, ×弱欺骗者  
Fig. 5 Strong Cors expt. ●weak C, ▲strong C, ×weak D

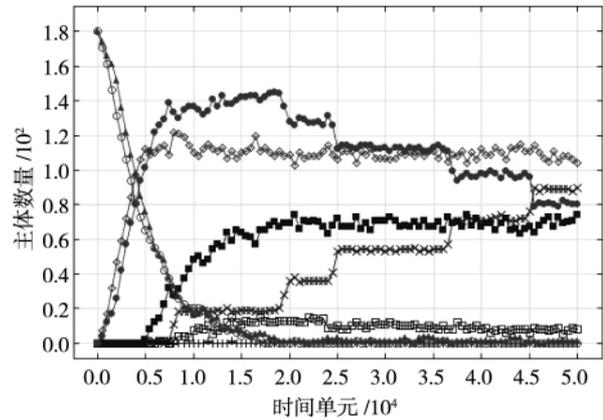


图6 强合作和强欺骗有条件联盟并存. 初始全部为强合作者.  
 $\eta = 80\%$ : ●弱合作者, ▲强合作者, ×弱欺骗者, □强欺骗者;  
 $\eta = 95\%$ : ◇弱合作者, ○强合作者, +弱欺骗者, ■强欺骗者  
Fig. 6 Strong Cors. and condi. union of strong Dors. coexist.  
 $\eta = 80\%$ : ●weak C, ▲strong C, ×weak D, □strong D;  
 $\eta = 95\%$ : ◇weak C, ○strong C, + weak D, ■strong D.

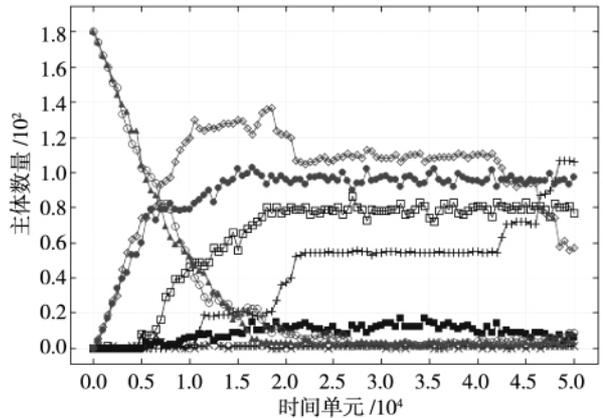


图7 强合作和“强欺骗双联盟”并存. 初始全部为强合作者  
 $\theta = 20\%$ : ●弱合作者, ▲强合作者, ×弱欺骗者, □强欺骗者  
 $\theta = 5\%$ : ◇弱合作者, ○强合作者, +弱欺骗者, ■强欺骗者  
Fig. 7 Strong Cors. and duo-union of strong Dors. coexist.  
 $\theta = 20\%$ : ●weak C, ▲strong C, ×weak D, □strong D;  
 $\theta = 5\%$ : ◇weak C, ○strong C, + weak D, ■strong D.

同时注意两个问题. 第 1, 可能存在一种“真诚”的利他主义局限. 强合作代表这种利他主义, 它并不一定能给社会带来最大福利, 反而可能造成福利损失. 这并非贬低利他精神, 但需对此保持谨慎. 第 2, 另一种“伪善”的利他主义可能给社会带来“灾难”. 强欺骗代表这种利他主义, 它对不合作的干预可能被误读为“利他”, 却可能导致社会陷入“伪善”的非合作困境. “伪善”是道德评价, 若将社会道德看作是社会成员合作生产的公共品, 强欺骗就是要求别人遵守道德而自己却不遵守的一种“伪善”.

进一步分析, 抛开道德评价, 如果将合作者看作是生产者提供公共产品, 那么惩罚者便是生产的监管者, 抑制生产中“搭便车”行为, 将不合作者“挤”出系统. 其中, 强合作者仍需要参加生产, 强欺骗者则无偿占有合作收益: 前者是“兼职”, 后者是“专职”. 上述混合策略均衡或可看作职能分工: 虽然监管者无偿占有公共品, 但也将保证生

产持续性. 这样, 最优水平的合作是不存在的, 但是可以在次优水平上获得稳定合作收益. 这时的社会产出是实际得到的最大化水平. 当然, 这个结果的前提是“监管者”中演化出了某种制衡机制, 从而削弱其剥削的能力. 制衡也并非完全可靠, 过多或过少的不稳定都会陷入非合作均衡.

## 4 结束语

理论上合作困境之所以在现实中被破解, 是因为现实中存在种种约束机制, 使人们选择合作. 无论是外生性约束, 还是内生性约束, 关键是由谁来提供约束成本, 谁来提供发展资源. 本文提供了形式化的解释, 即在合作与非合作的权衡中实现合作收益最大化. 但是本文在模型参数的合理性和系统的稳健性等方面还存在不足, 后面的工作将主要围绕这些展开.

### 参 考 文 献:

- [1] Hamilton W D. The genetical evolution of social behavior I/II [J]. *Journal of Theoretical Biology*, 1964, 7(1): 1-16.
- [2] Trivers R L. The evolution of reciprocal altruism [J]. *The Quarterly Review of Biology*, 1971, 46(1): 35-57.
- [3] Axelrod R. *The Evolution of Cooperation* [M]. New York: Basic Books, 1984.
- [4] Smith J M. *Evolution and the Theory of Games* [M]. Cambridge: Cambridge University Press, 1992.
- [5] Alexander R D. *The Biology of Moral Systems* [M]. Hawthorne: Aldine de Gruyter, 1987.
- [6] Zahavi A. Altruism as a handicap: The limitations of kin selection and reciprocity [J]. *Journal of Avian Biology*, 1995, 26(1): 1-3.
- [7] Nowak M A, Sigmund K. Evolution of indirect reciprocity by image scoring [J]. *Nature*, 1998, 393(6685): 573-577.
- [8] Nowak M A, Sigmund K. Evolution of indirect reciprocity [J]. *Nature*, 2005, 437(7063): 1291-1298.
- [9] Brandt H, Sigmund K. Indirect reciprocity, image scoring, and moral hazard [J]. *Proc Natl Acad Sci USA*, 2005, 102(7): 2666-2670.
- [10] Lotem A, Fishman M A, Stone L. Evolution of cooperation between individuals [J]. *Nature*, 1999, 400(6741): 226-227.
- [11] Fishman M A. Indirect reciprocity among imperfect individuals [J]. *Journal of Theoretical Biology*, 2003, 225(3): 285-292.
- [12] Roberts G. Evolution of direct and indirect reciprocity [J]. *Proceedings of the Royal Society B*, 2008, 275(1631): 173-179.
- [13] Fehr E, Gächter S. Altruistic punishment in humans [J]. *Nature*, 2002, 415(6868): 137-140.
- [14] Gintis H. Strong reciprocity and human sociality [J]. *Journal of Theoretical Biology*, 2000, 206(2): 169-179.
- [15] Bowles S, Gintis H. The evolution of strong reciprocity: Cooperation in heterogeneous populations [J]. *Theoretical Population Biology*, 2004, 65(1): 17-28.
- [16] Quervain D J-F de, Fischbacher U, Treyer V, et al. The neural basis of altruistic punishment [J]. *Science*, 2004, 305(1254): 1254-1258.
- [17] Axelrod R. An evolutionary approach to norms [J]. *The American Political Science Review*, 1986, 80(4): 1095-1111.
- [18] Kameda T, Takezawa M, Hastie R. The logic of social sharing: An evolutionary game analysis of adaptive norm develop-

- ment[J]. *Personality and Social Psychology Review*, 2003, 7(1): 2–19.
- [19] Janssen M A, Ostrom E. Adoption of a new regulation for the governance of common-pool resources by a heterogeneous population[C]// Baland J M, Bardhan P, Bowles S. eds. *Inequality, Cooperation and Environmental Sustainability*. NJ: Princeton University Press, 2006: 60–96.
- [20] Henrich J. Cultural group selection, coevolutionary processes and large-scale cooperation[J]. *Journal of Economic Behavior & Organization*, 2004, 53(1): 3–35.
- [21] Henrich J, Boyd R. Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas[J]. *Journal of Theoretical Biology*, 2001, 208(1): 79–89.
- [22] Tummolini L, Castelfranchi C, Henrich J, et al. Culture, evolution and the puzzle of human cooperation[J]. *Cognitive Systems Research*, 2006, 7(2): 220–245.
- [23] Boyd R, Gintis H, Bowles S, et al. The evolution of altruistic punishment[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2003, 100(6): 3531–3535.
- [24] Bowles S, Choi Jung-Kyoo, Hopfensitz A. The co-evolution of individual behaviours and social institutions[J]. *Journal of Theoretical Biology*, 2003, 223(2): 135–147.
- [25] Boyd R, Richerson P J. Punishment allows the evolution of cooperation (or anything else) in sizable groups[J]. *Ethology and Sociobiology*, 1992, 13(3): 171–195.
- [26] Kiyonari T, Shimoma E, Yamagishi T. Second-order punishment in one-shot social dilemma[J]. *International Journal of Psychology*, 2004, 39(5/6): 329.
- [27] Andreoni J, Harbaugh W, Vesterlund L. The carrot or the stick: Rewards, punishments, and cooperation[J]. *American Economic Review*, 2003, 93(3): 893–902.
- [28] Oliver P. Rewards and punishments as selective incentives for collective action: Theoretical investigations[J]. *The American Journal of Sociology*, 1980, 85(6): 1356–1375.
- [29] Oliver P. Rewards and punishments as selective incentives: An apex game[J]. *The Journal of Conflict Resolution*, 1984, 28(1): 123–148.
- [30] Walker J M, Halloran M A. Rewards and sanctions and the provision of public goods in one-shot settings[J]. *Experimental Economics*, 2004, 7(3): 235–247.
- [31] Sefton M, Shupp R, Walker J M. The effect of rewards and sanctions in provision of public goods[J]. *Economic Inquiry*, 2007, 45(4): 671–690.
- [32] Sigmund K, Hauert C, Nowak M A. Reward and punishment[J]. *Proceedings of the National Academy Sciences of the United States of America*, 2001, 98(19): 10757–10761.
- [33] David G R, Anna D, Tore E, et al. Positive interactions promote public cooperation[J]. *Science*, 2009, 325(5945): 1272–1275.
- [34] Milinski M, Semmann D, Krambeck H. Reputation helps solve the ‘tragedy of the commons’ [J]. *Nature*, 2002, 415(6870): 424–426.
- [35] Panchanathan K, Boyd R. Indirect reciprocity can stabilize cooperation without the second-order free rider problem[J]. *Nature*, 2004, 432(7016): 499–502.
- [36] 张 路, 王浣尘, 陈 忠. 合作进化的研究进展[J]. *上海交通大学学报*, 1999, 33(10): 1322–1328.  
Zhang Lu, Wang Huanchen, Chen Zhong. Review on the evolution of cooperation[J]. *Journal of Shanghai Jiaotong University*, 1999, 33(10): 1322–1328. (in Chinese)
- [37] 叶 航, 汪丁丁, 罗卫东. 作为内生偏好的利他行为及其经济学意义[J]. *经济研究*, 2005, (8): 84–94.  
Ye Hang, Wang Dingding, Luo Weidong. On endogenetic preference for altruism behaviors and its implications to economics[J]. *Economic Research Journal*, 2005, (8): 84–94. (in Chinese)
- [38] 韦 森. 从合作的演化到合作的复杂性: 评阿克罗斯德关于人类合作生成机制的博弈论试验及其相关研究[J]. *东岳论丛*, 2007, 28(3): 1–20.  
Wei Sen. From the evolution of cooperation to the complexity of cooperation: Commentary on the Axelrod’s game test and relevant studies about the generation mechanism of human cooperation[J]. *Dong Yue Tribune*, 2007, 28(3): 1–20. (in Chinese)
- [39] 韦 倩. 强互惠理论研究评述[J]. *经济学动态*, 2010, (5): 106–111.  
Wei Qian. Review on the research of strong reciprocity theory[J]. *Economic Perspectives*, 2010, (5): 106–111. (in Chinese)

- [40]黄 璜. 合作进化模型综述[J]. 北京大学学报(自然科学版), 2011, 47(1): 185–192.  
Huang Huang. Review on model of cooperation evolution[J]. Acta Scientiarum Naturalium Universitatis Pekinensis, 2011, 47(1): 185–192. (in Chinese)
- [41]石 磊, 王瑞武. 合作行为的非对称性演化[J]. 中国科学: 生命科学, 2010, 40(1): 62–72.  
Shi Lei, Wang Ruiwu. Asymmetrical evolution of cooperation behavior[J]. Scientia Sinica Vitae, 2010, 40(1): 62–72. (in Chinese)
- [42]黄 璜. 基于社会资本的合作演化研究——“基于主体建模”方法的博弈推演[J]. 中国软科学, 2010, (9): 173–184.  
Huang Huang. Cooperative evolution based on social capital: A simulation using agent-based modeling[J]. China Soft Science, 2010, (9): 173–184. (in Chinese)
- [43]蒋国银, 胡 斌. 集成博弈和多智能体的人群工作互动行为研究[J]. 管理科学学报, 2011, 14(2): 29–41.  
Jiang Guoyin, Hu Bin. Study on interaction behavior between group and work based on game and multi-agent[J]. Journal of Management Sciences in China, 2011, 14(2): 29–41. (in Chinese)
- [44]张洪恩, 王覃刚. 强互惠理论的扩展[J]. 中国工业经济, 2007, (3): 70–78.  
Zhang Hong'en, Wang Qingang. The extension of strong reciprocity theory[J]. China Industrial Economy, 2007, (3): 70–78. (in Chinese)
- [45]王飞跃, 史帝夫·兰森. 从人工生命到人工社会——复杂社会系统研究的现状和展望[J]. 复杂系统与复杂性科学, 2004, 1(1): 33–41.  
Wang Feiyue, Lansing J S. From artificial life to artificial societies: New methods for studies of complex social systems[J]. Complex Systems and complexity Science, 2004, 1(1): 33–41. (in Chinese)
- [46]黄 璜. 社会科学研究中“基于主体建模”方法评述[J]. 国外社会科学, 2010, (5): 40–47.  
Huang Huang. Review on agent-based modeling in social science research[J]. Social Sciences Abroad, 2010, (5): 40–47. (in Chinese)
- [47]Heckathorn D D. Collective action and the second-order free-rider problem[J]. Rationality and Society, 1989, 1(1): 78–100.
- [48]Oliver P, Marwell G, Teixeira R. A theory of the critical mass. I. Interdependence, group heterogeneity, and the production of collective action[J]. The American Journal of Sociology, 1985, 91(3): 522–556.
- [49]Kameda T, Takezawa M, Hastie R. The logic of social sharing: An evolutionary game analysis of adaptive norm development[J]. Personality and Social Psychology Review, 2003, 7(1): 2–19.

## Logic of cooperation: An evolutionary analysis of strong defection strategy

*HUANG Huang*

School of Government, Peking University, Beijing 100871, China

**Abstract:** How a high level of cooperation emerges in human societies, especially large-scale societies, in spite of a low level of genetic relatedness among society members? In this paper, based on the N-player Iterated Prisoner's Dilemma model, we offer a model of evolution of cooperation based on punishment using agent-based modeling approach. We introduce two different types of punishment strategies in our model. One is the strong cooperation strategy which means cooperating and punishing those defecting. The other is called the strong defection strategy where one defects and punishes others who defect. Results of model simulation show that the strong cooperator will hinder cooperation in the macro level though it is good in the micro level, and the strong defection, a non-cooperation strategy indeed, may facilitate social cooperation even in a large-scale society, because when those agents who are strong defectors are conditionally united or form in multiple unions, a cooperation-defection mixed-strategy equilibrium will emerge, which means that there is a sub-optimal level of cooperation in social systems though not an optimal one.

**Key words:** evolution of cooperation; punishment; strong defection; unstable union; large-scale social cooperation