结构化种群中促进合作的制度激励优化

Shengxian Wang,[1,2] Xiaojie Chen,[1,*] Zhilong Xiao,[1,3] Attila Szolnoki[4] 和 V´ıtor V. Vasconcelos[5,6]

[1]中国电子科技大学数学科学学院，成都 *611731*，中国

[2]格罗宁根大学理工学院，荷兰格罗宁根 *9747 AG*[3]中山大学计算机科学与工程学院，中国广州 *510006*[4]技术物理与材料科学研究所，能源研究中心，邮政信箱 *49*，匈牙利布达佩斯 *H-1525*[5]阿姆斯特丹大学信息学研究所计算科学实验室，荷兰阿姆斯特丹 *1098XH*[6]阿姆斯特丹大学高等研究院，荷兰阿姆斯特丹 *1012 GC*

arXiv:2301.06472v1 [cs.GT] 16 Jan 2023

奖励和惩罚等激励措施，是促进结构化种群中互动个体之间合作的一种常用方式。然而，如何恰当地运用激励，仍是激励提供机构面临的一项挑战。特别是，鉴于实施激励需要付出成本，探索能够以最低成本实现预期集体目标的最优激励方案，值得研究。在本研究中，我们分别考察了正向和负向激励，研究对象是互动冲突由囚徒困境博弈描述的结构化个体种群。我们建立了一个指标函数，用于量化激励实施过程中的累积成本，并从理论上推导出规则网络上促进合作的最优正向和负向激励方案。我们发现，这两类最优激励方案完全相同，且不随时间变化。此外，我们比较了奖励方案和惩罚方案在实施成本方面的差异，并为在博弈论框架中使用激励提供了严格依据。我们还进行了计算机模拟，以支持理论结果并探究其在不同种群结构中的稳健性，包括规则网络、随机网络、小世界网络和无标度网络。

健性，包括规则网络、随机网络、小世界网络和无标度网络。

*电子邮箱：xiaojiechen@uestc.edu.cn

1. 引言

在当今时代，合作至关重要 [1]。然而，合作的演化与涌现却与互动个体的即时自利利益相冲突 [2]。演化博弈论提供了一个通用的数学框架，用以描绘个体如何在动态过程中相互作用，并预测合作行为如何在种群层面演化 [3]。作为一种代表性范式，囚徒困境博弈在研究互动个体种群中的合作问题方面受到了广泛关注 [4]。

当个体在一个大型、充分混合的种群中进行演化囚徒困境博弈时，由于所有个体相互互动的概率相同，合作无法涌现。然而，现实世界中的种群结构并非充分混合，而是相对复杂的；在结构化种群中，个体之间的互动仅限于一组邻居 [5]。受网络科学快速发展的启发，研究者检验了若干交互拓扑，包括小世界网络 [6] 和无标度网络 [7]。这些网络的特征会显著影响合作的演化动力学 [8–14]。除网络拓扑之外，合作演化的另一个关键决定因素是策略更新规则，它决定了微观层面的更新过程 [15–18]。例如，如果囚徒困境博弈中的收益成本比超过特定交互网络的平均度，那么在所谓的死亡-出生策略更新规则下，合作可能受到促进 [15–17]。相比之下，在另一种出生-死亡策略更新规则下，合作则无法涌现 [15]。

总的来说，在未纳入额外调节机制或道德助推的情况下，结构化种群中有许多情形无法涌现合作 [15, 19, 20]。在这些不利环境中，可以利用亲社会激励，在彼此无亲缘关系且相互竞争的个体之间维持合作 [21–29]。具体而言，可以奖励合作者的积极行为，或惩罚背叛者造成的广泛负面影响。在个体层面，只要激励的数额超过合作与背叛之间的收益差，无论激励是正向还是负向，都能促进合作。然而，在不利于合作涌现的特定环境中，究竟需要多强的正激励或负激励才能推动种群朝期望的方向发展，仍不清楚。此外，实施激励总是需要付出成本 [30–34]。以往相关研究假设，激励一旦实施，其数额便固定在某个值 [30, 32]。在这种情况下，所获得的激励方案未必具有最低的累积成本 [33]。因此，寻找最优的时变激励方案，对于确保干预有效、推动

驱动种群 种群以最低的实施成本迈向富有成效且合作的状态，至关重要。

我们的研究探讨合作涌现所需的激励力度，并在博弈论框架下研究最优激励方案：为参与囚徒困境博弈的结构化种群中的个体提供随时间变化的机构正向或负向激励。我们建立了一个指标函数，用于量化执行过程中的累积成本。我们系统地考察了所有相关的策略更新规则，包括死亡-出生（DB）、出生-死亡（BD）、模仿（IM）和成对比较（PC）更新 [15–17, 35]。利用最优控制理论，我们针对每种策略更新规则，求得了使合作涌现的累积成本最小化的动态激励方案。有趣的是，我们发现，对于给定的策略更新规则，最优负向激励方案与最优正向激励方案相同，且均不随时间变化。然而，如果初始合作水平大于完全合作状态与期望合作状态之差，施加惩罚比给予奖励所需的累积成本更低；否则，给予奖励所需的成本更低。除解析计算外，我们还进行了计算机模拟，证实我们的结果适用于广泛的种群结构。

2. 结果

我们从一个由 N 个体组成的结构化种群出发，这些个体在度为$k > 2$的正则网络中相互作用。在该图中，顶点代表相互作用的个体，边则决定个体之间的互动关系。每个个体 i 都与其邻居进行囚徒困境博弈。个体可以选择成为合作者（C），向对手提供收益 b，自身则付出成本 c；也可以选择成为背叛者（D），无需付出成本，也不向他人提供任何收益。与一个邻居进行博弈后，如果实施了制度性正向激励，该个体选择 C便可获得金额为μ_R的激励奖励。如果实施了制度性负激励，该个体选择 D便会被处以金额为 μ_P的罚款。该个体通过与所有邻居互动，累计获得收益 π_i。我们将个体 i的适应度 f_i（主要指繁殖率）设定为 $1-\omega+\omega\pi_i$，其中 ω （$0 \leq \omega \leq 1$）衡量选择强度 [36]。本文主要研究弱选择的影响，即 $0 < \omega \ll 1$。我们分别考察四种依赖适应度的策略更新规则，详情见模型与方法部分。此外，为了帮助读者直观理解网络化囚徒困境博弈中实施制度奖励或惩罚时的演化过程，我们给出了示意图，如图1所示。

2.1 最优激励方案的理论预测

在本小节中，我们分别探讨四种不同策略更新规则下的激励方案，包括 DB、BD、模仿者和 PC 更新。这些计算的详细过程见电子补充材料；此处我们仅概述弱选择极限下，具有正向或负向激励的动力系统所采用的成对近似方法。

采用对偶近似方法（见电子补充材料第 1 节），我们得到 DB 规则下实施正向或负激励时合作者比例的动力学方程：

$$\frac{dp_C}{dt} = F_{\mathrm{DB}}(p_C, \mu_v, t) = \frac{\omega(k-2)}{k-1}[b + k(\mu_v - c)]p_C(1-p_C) + o(\omega^2)\,, \tag{1}$$

其中，p_C 是整个种群中的合作者比例；μ_v 是合作者从激励提供机构获得的正向激励量，或中央机构施加于背叛者的负激励量（请注意，如模型与方法部分所述，每次互动都会施加此类激励）。方程 (1) 有两个均衡点，一个位于 $p_C = 0$，另一个位于 $p_C = 1$。如果 $\mu_v > c - \frac{b}{k}$，则第一个均衡点不稳定，第二个稳定，这表明合作者胜过背叛者（更多细节见电子补充材料第 1 节）。值得注意的是，在没有激励的情况下，即 $\mu_v = 0$，我们重新得到先前确定的合作演化条件 $b/c > k$ [15, 16]。

利用 Hamilton-Jacobi-Bellman（HJB）方程 [37–39], 在 $\mu_v > c - \frac{b}{k} > 0$ 的条件下，我们解析地得到了奖励和惩罚的最优方案 $\mu_v^* = \frac{2}{(ck-b)\,k}$（详情见电子补充材料第 1 节）。最优奖励和最优惩罚方案均与时间无关，且惩罚和奖励的最优激励水平相同，即 $\mu_v^* = \mu_R^* = \mu_P^*$。因此，采用最优奖励或最优惩罚方案，可以求解方程（1）所描述的动力系统，其解为$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\mathrm{DB}}t}}$，其中 $\beta_{\mathrm{DB}} = \frac{\omega\,(k-2)}{(ck-b)\,k-1}$，而 $p_0 = p_C(0) > 0$表示种群中的合作者初始比例。

奖励方案 μ_R^* 使动力系统从初始状态 p_0到达预期终止状态 $p_C(t_f)$所产生的累积成本为

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R^*)^2}{2}dt = \frac{(kN\mu_R^*)^2}{2\beta_{\mathrm{DB}}}[p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})], \tag{2}$$

类似地，惩罚协议 μ_P^* 所产生的累积成本为

$$J_P^* = \int_0^{t_f} \frac{(kNp_D\mu_P^*)^2}{2} dt = \frac{(kN\mu_P^*)^2}{2\beta_{\text{DB}}}[p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]. \tag{3}$$

我们进一步发现，$J_R^* > J_P^*$若 $p_0 > \delta$，且$J_R^* < J_P^*$若 $p_0 < \delta$（详情请参阅电子补充材料第 1 节）。

对于 BD 更新，动力学方程为

$$\frac{dp_C}{dt} = F_{\text{BD}}(p_C, \mu_v, t) = \frac{\omega k(k-2)}{k-1}(\mu_v - c)p_C(1-p_C) + o(\omega^2). \tag{4}$$

我们证明，当 $\mu_v > c$时，系统能够达到稳定的完全合作状态。自然地，在没有激励的情况下，我们会得到 Ohtsuki等人 [15, 16]的结果。通过求解 HJB方程，奖励和惩罚的最优激励方案均为 $\mu_v^* = 2c$。方程 (4) 的解为$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{BD}}t}}$，其中 $\beta_{\text{BD}} = {}^{\omega k}{}_{(k-2)_{k-1}}{}^{c}$（参见电子补充材料第 2 节）。因此，在最优奖励方案下达到预期终止状态的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{BD}}}[p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})], \tag{5}$$

而对于最优惩罚方案，则为

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{BD}}}[p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]. \tag{6}$$

在 IM 更新的情况下，动力学方程变为

$$\frac{dp_C}{dt} = F_{\text{IM}}(p_C, \mu_v, t) = \frac{\omega k^2(k-2)}{(k+1)^2(k-1)}[b + (\mu_v - c)(k+2)]p_C(1-p_C) + o(\omega^2)\,, \tag{7}$$

这表明，当 $\mu_v > c - \frac{b}{k+2}$ 时，系统会演化至完全合作状态。最优激励方案的HJB 方程解给出 $\mu_v^* = {}^{2[c}{}_{(k+2\ -b]\ )}{}^{k+2}$，奖励和惩罚两种情形均如此（见电子补充材料第 3 节）。方程 (7) 的解为 $p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{IM}}t}}$，其中 $\beta_{\text{IM}} = {}^{\omega k^2}{}_{(k-2)[c(k+2)-b]\ (k+1)^2(k-1)}$。最优奖励方案所需的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{IM}}}[p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})], \tag{8}$$

而对于最优惩罚方案，则变为

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{IM}}}[p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})]. \tag{9}$$

采用 PC 更新时，动力学方程为

$$\frac{dp_C}{dt} = F_{\text{PC}}(p_C, \mu_v, t) = \frac{\omega k(k-2)}{2(k-1)}(\mu_v - c)p_C(1-p_C) + o(\omega^2), \tag{10}$$

由此可知，当满足 $\mu_v > c$ 时，可以达到稳定的完全合作状态（见电子补充材料第 4 节）。对于最优激励方案，我们得到 $\mu_v^* = \mu_R^* = \mu_P^* = 2c$。方程（10）的解为 $p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{PC}}t}}$，其中 $\beta_{\text{PC}} = \frac{\omega k}{(k-2)2\ k-1}()$。因此，最优奖励方案的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{PC}}}[p_0 + \delta - 1 + \ln(\frac{1-p_0}{\delta})], \tag{11}$$

而

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{PC}}}[p_0 + \delta - 1 + \ln(\frac{1-\delta}{p_0})], \tag{12}$$

则是最优惩罚方案的累积成本。

请注意，对于每种更新规则，弱选择极限下的控制方程总有两个平衡点，分别对应完全背叛状态和完全合作状态。因此，我们可以求得合作演化所需激励的最小量，如图 2 所示。无论采用哪种更新规则，最优激励方案都是随时间不变的，且奖励与惩罚所需的方案相同。因此，对于每种更新规则，$\mu_v^* = \mu_R^* = \mu_P^*$。最优方案汇总于图 2。我们还在图 2 中展示了每种更新规则下最优奖励与惩罚方案的累积成本。

2.2. 最优激励方案的数值计算与计算机模拟

下文将通过数值计算和计算机模拟，验证我们关于 $\mu_v^*(v \in \{R, P\})$ 的解析结果。图 3 展示了在我们考察的四种策略更新规则下，合作者比例 p_C 随时间变化的情况 t，其中比较了最优 μ_v^* 和另外两种激励方案。首先需要指出的是，对于每种更新规则，从随机的 $p_0 = 0.5$ 状态开始演化时，系统总会到达预期终止状态 $p_C(t_f) = 0.99$。然而，达到这一状态所需的激励提供机构成本存在显著差异；无论采用奖励还是惩罚，μ_v^* 的成本都是最低的。值得注意的是，最优激励方案并不会带来最快的弛豫。

此外，我们展示了针对每种更新规则施加正向或负向激励时的蒙特卡洛模拟结果（电子补充材料，图 S1-4）。我们发现，对于不同的更新规则，采用最优激励方案并不会使系统最快达到

期望的合作状态，但始终能使机构的累积成本最低。我们必须强调，我们的发现并不局限于正则网络，而是适用于广泛的交互图，从不规则随机网络到小世界网络和无标度网络均然。我们的数值计算和模拟结果与解析预测一致，并表明采用最优激励方案时累积成本最低。

2.3. 最优奖励与惩罚方案的比较。

通过电子补充材料中给出的理论分析，我们可以得出结论：对于每种更新规则，如果初始合作水平大于完全合作状态与期望合作状态之间的差值，则与最优奖励方案相比，执行最优惩罚方案所需的累积成本更低；反之，采用最优奖励方案所需的累积成本更低。为验证这一理论预测并进行直观比较，我们展示了初始合作者比例 p_0 可调时，最优惩罚与奖励方案所产生的累积成本值。假设采用最优激励方案，所需的累积成本可通过数值积分或蒙特卡洛模拟确定。

针对每种更新规则，我们的数值结果汇总于图 4。由图 4 上排可见，当初始合作水平大于完全合作状态与期望终止状态之间的差值（对背叛设定容差 δ）时，机构通过惩罚达到预期合作状态所需的累积成本较低。相反，由图 4 下排可见，当初始合作水平小于完全合作状态与期望终止状态之间的差值时，机构通过奖励达到预期合作状态所需的累积成本较低。然而，在这些不同情形下，无论采用最优奖励方案还是最优惩罚方案，相应的累积成本值都会随着初始合作水平 p_0的增加而降低。此外，电子补充材料中图 S5 所示的模拟结果也支持我们的理论分析，为在结构化种群的演化囚徒困境博弈中，根据不同初始条件使用激励提供了严谨的理论依据。

3. 讨论

在人类社会中，亲社会激励是避免“公地悲剧

”的必要手段之一 [21, 22, 40]。然而，对于激励提供机构而言，激励的选择主要基于两个方面。其一是了解在结构化种群中需要多大的激励才能促进合作演化（可能适用于不同的更新规则）。其二是所采用的激励方案是否为一种能使机构成本最低的最优时变协议。

为研究上述问题，我们建立了一个博弈论框架。具体而言，我们分别考察在采用四种不同策略更新规则的网络化囚徒困境博弈中引入正向或负激励。对于给定的更新规则，我们得出了合作演化所需的最低激励量的理论条件。通过建立一个用于量化执行成本的指标函数，我们运用HJB方程的方法，分别推导出每种策略更新规则下的最优正向和负激励方案。我们发现，对于所有考察的更新规则，这些最优激励方案均不随时间变化。此外，负激励与正向激励对应的最优激励方案完全相同。然而，当初始合作水平相对较高时，采用惩罚方案可使激励提供机构的累积成本更低；否则，采用奖励方案的成本更低。我们还进行了计算机模拟，结果证实，我们的结果适用于由规则网络、随机网络、小世界网络和无标度网络所描述的不同交互拓扑，从而表明这些发现具有普遍的稳健性。

在本研究中，我们量化了在所考虑的每种更新规则下，合作演化所需的激励强度。然而，提供亲社会激励时，博弈结构可能会发生变化。具体而言，当激励金额为 $\mu_v > c$ 时，囚徒困境博弈将转变为和谐博弈，此时合作者会自然地胜过背叛者 [9]。有趣的是，我们注意到，在 DB 更新下，如果 $\mu_v > c-b/k$；在 IM 更新下，如果 $\mu_v > c-b/$（$k+2$），合作就会受到促进。这意味着，在 DB 和 IM 规则下，μ_v 值可以小于成本 c，从而确保博弈结构不变。然而，对于 BD 和 PC 更新，合作演化所需的激励金额必须完全超过合作成本。因此，干预的有效性或许能为哪种更新类型最能反映种群行为提供一些线索。

在此，我们强调，通过求解所构建的最优控制问题，所得的最优激励方案具有时间不变性。更值得注意的是，对于每种更新规则，所得的最优负向激励协议与最优正向激励协议相同。因此，这些

最优激励方案与状态无关。因此，机构无需时刻监测种群状态即可实施最优激励，从而节省监测成本，因为监测通常成本高昂。此外，针对数据库更新和模仿者更新的最优激励水平能够确保个体面临的困境仍然是囚徒困境。因此，我们的研究表明，在数据库更新和模仿者更新下获得的基于激励的控制方案简单有效，有助于促进结构化种群中的合作演化。

激励可作为调控工具，用于规范个体的决策行为 [28]。然而，从个体和激励提供机构的角度来看，在合作演化过程中，对惩罚和奖励的使用存在显著的偏好差异 [41]。有研究指出，人们往往不倾向于惩罚，因为在重复博弈中，使用惩罚会导致总收入或平均收益偏低[23, 24]。相比之下，激励提供机构更倾向于频繁使用惩罚，这是因为惩罚能以较低的激励实施成本来促进合作 [30, 41]。本文考察一种自上而下式激励机制：在该机制下，合作者可获得奖励，或背叛者可由一个已存在且稳定运作的外部集中式机构直接惩罚。在这一框架中，我们严格比较奖励方案和惩罚方案的实施成本，并阐明最佳选择如何取决于初始合作水平，从而为在囚徒困境博弈中使用激励及其在人类社会中的潜在应用提供严谨依据。

我们的工作还有许多可拓展之处，尤其是在解析方面。我们想指出，我们采用成对近似方法获得解析结果，而这种方法主要用于规则网络 [15]。事实上，只要进一步考虑这些网络的特性，这种方法也可以拓展到其他类型的复杂网络 [42, 43]。沿着这一思路，值得在充分考虑底层种群结构空间特性的基础上，从解析角度研究促进合作演化的低成本激励策略。此外，我们还可以采用基于计算随机游走合并时间的方法，拓展理论分析 [11]。当需要在拓扑层面进行干预时，这些方法尤其有用，因为它们能更紧密地将结果与网络联系起来。此外，我们是在弱选择极限下得到理论结果的。其他情境下的数值分析表明，受环境相关因素的影响，无论种群

是充分混合还是具有结构，选择强度都会改变博弈动态，并在决定高成本效益的制度激励 [44–47]方面发挥关键作用。因此，一个自然的问题是：当选择并非弱选择时，我们的理论结果是否仍然成立 [48]。事实上，对于某些具有高对称性的结构，强选择下的解析计算仍然可行。因此，我们有望确定促进合作所需激励程度的理论条件，并探索适用于这些种群结构的最优激励方案。

此外，我们的研究旨在将种群中的合作水平提升至设定值，同时尽可能降低激励成本，而不考虑达到该水平需要多长时间。对于转变速率可能至关重要的情形，仍需进一步研究；时变方案或许可以解决这一问题。我们在囚徒困境博弈框架下开展研究，该博弈是研究合作演化的典范。还有其他典型的双人困境，例如雪堆博弈 [49] 和鹿猎博弈 [50]。未来研究可将本研究拓展至上述博弈，并考察群体互动或高阶互动 [51–55]。此外，我们通过考虑全局信息（即整个种群中的合作者比例）来设计成本效益高的激励方案，但结构化网络上的局部邻域属性，例如邻域中有多少合作者，会影响最终的演化结果 [56–58]。因此，未来研究应将这些局部信息纳入考量，以设计成本最低的最优激励方案。最后，诚实 [59] 或信任与可信赖性 [60], 等其他亲社会行为对合作也至关重要，因此，研究用于促进这些行为演化的激励优化问题，是一项有意义的拓展。

4. 模型与方法

4.1. 囚徒困境博弈

我们考虑将一群个体分布在交互图的各个节点上。在每一轮中，每个个体都与其邻居进行演化囚徒困境博弈，并可以选择合作（C）或背叛（D）。我们将该博弈的收益矩阵设为

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array}, \tag{13}$$

其中，b 表示合作收益，c（$0 < c < b$）表示合作成本。在与所有相邻邻居进行成对交互后，每个个体根据收益矩阵获得其收益。

4.2. 制度激励

此外，激励提供机构提供的亲社会激励可用于在合作者与邻居进行博弈后奖励合作者或惩罚背叛者。在此，我们分别考虑两种激励类型，即正向激励和负向激励。若采用正向激励，则博弈中的合作者与邻居 [24, 28] 互动时，可从中央机构获得金额为 μ_R 的奖励。若采用负向激励，则背叛者每次互动 [23] 都会被罚款 μ_P。因此，采用正向激励时，修正后的收益矩阵为

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & \quad D \end{array} \\ \begin{pmatrix} b-c+\mu_R & -c+\mu_R \\ b & 0 \end{pmatrix} \end{array}, \tag{14}$$

采用负向激励时，修正后的收益矩阵为

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & \quad D \end{array} \\ \begin{pmatrix} b-c & -c \\ b-\mu_P & -\mu_P \end{pmatrix} \end{array}. \tag{15}$$

因此，根据上述收益矩阵，每个个体都会收集其总收益，该收益来自与邻居及激励提供机构的成对互动。

4.3. 策略更新规则

根据进化选择原则，玩家会不时更新策略，但更新方式可能会显著影响进化结果 [9, 15]。与先前的研究 [15, 16], 一致，我们在此考虑四种主要的策略更新规则，分别描述 DB 更新、BD 更新、模仿者更新和个人计算机更新。具体而言，在 DB 更新中，每个时间步都会从整个种群中随机选取一个个体使其死亡；随后，邻居以与其适应度成正比的概率竞争这个空位。在 BD 更新中，每个时间步都会以与适应度成正比的概率从整个种群中选取一个个体进行繁殖；该个体的后代会取代一个随机选中的邻居。在模仿者更新中，每个时间步都会从整个种群中随机选取一个个体来更新其策略；它要么保留自己的策略，要么以

与其适应度成正比的概率模仿某个邻居的策略。在个人计算机更新中，每个时间步都会随机选取一个个体来更新其策略，并将自身适应度与一个随机选取的邻居进行比较。焦点个体要么保留当前策略，要么以取决于适应度差异的概率采纳邻居的策略。

4.4. 优化激励

由于提供激励会给机构带来成本，因此找到一种最优的 μ_v^* 至关重要：它既能以最小的投入有效支持合作，又能达到预期效果。为此，我们首先建立一个指标函数，用于量化执行过程中的累积成本，其表达式为

$$J_v = \int_{t_0}^{t_f} \frac{(kNp_i\mu_v)^2}{2} dt, \tag{16}$$

其中，若 $p_i = p_C$，则 $v = R$，这意味着实施正向激励；否则，$p_i = p_D$，这意味着实施负向激励。这里，t_0 为初始时间，在本文中设为 0；t_f 为系统的终止时间。基于上述描述，我们接下来运用最优控制理论 [37–39]，探究从 0 到 t_f 这一演化时期内的最优激励方案。一个关键假设是，t_f 并非固定不变；我们会持续监测演化过程，直至合作者比例在 t_f达到目标水平 $p_C(t_f)$。这里，我们假设 $p_C(t_f) = 1 - \delta > p_0$，其中 $p_0 (p_0 > 0)$ 为初始合作水平，而 δ是决定 t_f预期合作水平的参数，且满足 $0 \leq \delta < 1 - p_0$。

因此，我们将奖励或惩罚的最优控制问题表述如下：

$$\begin{aligned} \min \quad & J_v = \int_0^{t_f} \frac{(kNp_i\mu_v)^2}{2} dt, \\ s.t. \quad & \begin{cases} \frac{dp_C}{dt} = F_j(p_C, \mu_v, t), j = \text{DB, BD, IM, or PC}, \\ p_C(0) = p_0, \\ p_C(t_f) = 1 - \delta. \end{cases} \end{aligned} \tag{17}$$

这里，成本函数 J_v刻画了动力系统在时间段 $[0, t_f]$内从初始状态 p_0达到终止状态 $1-\delta$ 的平均累积成本。因此，数量 $\min J_v$ 可以作为计算具有最小执行成本的最优激励方案 μ_v^* 的目标。求解最优控制问题的详细信息见电子补充材料第 1—4 节。

4.5. 蒙特卡洛模拟

在一个完整的蒙特卡罗步中，平均而言，每个玩家都有机会更新其策略。上文已说明所采用的四种不同更新规则。此外，我们还测试了其他交互拓扑，包括使用大小为 $N = L \times L$ [4] 的二维方格晶格生成的正则网络，以及使用优先连接模型 [7] 获得的无标度网络；该模型从 $m_0 = 6$ 开始，在每个时间步中，每个新节点都与 $m = 2$ 个满足标准幂律分布的已有节点相连。作为替代方案，我们还考虑了 Erd˝os-R´enyi 随机图模型 [5]，以及采用 Watts-Strogatz 模型、具有 $p = 0.1$ 重连参数 [6] 的小世界网络。这些模拟结果汇总于电子补充材料的图 S1—S5 中。

作者贡献

X.C.、S.W.、V.V.V. 和 A.S. 设计了研究，S.W. 和 Z.X. 开展了研究，X.C.、V.V.V.、A.S. 和 S.W. 撰写了稿件，所有作者均讨论了研究结果，并对稿件提出意见和加以完善。

数据可获取性

本文没有补充数据。

利益冲突

作者声明不存在相互竞争的财务利益。

伦理

本文未报告涉及伦理考量的研究。

资金支持

本研究得到国家自然科学基金委员会（批准号：61976048和62036002）以及中国中央高校基本科研业务费的资助。S.W.感谢国家留学基金管理委员会（批准号：202006070122）的支持。A.S.得到国家研究、开发与创新办公室（NKFIH）项目K142948的资助。V.V.V.感谢阿姆斯特丹大学信息学研究所计算科学实验室提供的资助。

参考文献

[1] Hauert S、Mitri S、Keller L、Floreano D。2010年，《演化合作：从生物学到工程学》。进化机器人学的前沿。美国波士顿：麻省理工学院出版社。[2] Perc M、Jordan J J、Rand D G、Wang Z、Boccaletti S、Szolnoki A。2017年，人类合作的统计物理学。《物理学报告》687，1-51。（doi:10.1016/j.physrep.2017.05.004）[3] Hofbauer J、Sigmund K。1998年，《演化博弈与种群动力学》。英国剑桥：剑桥大学出版社。[4] Nowak M A、May R M。1992年，演化博弈与空间混沌。《自然》359，826-829。（doi: 10.1038/359826a0）[5] Erd˝os P、R´enyi A。1959年，关于随机图 I。《德布勒森数学出版物》6，290-297。（doi: 10.1109/IC-SMC.2006.384625）[6] Watts D J、Strogatz S H。1998年，“小世界”网络的集体动力学。《自然》393，440-442（doi: 10.1038/30918）。[7] Barab´asi A L、Albert R。1999年，随机网络中的标度涌现。《科学》286，509-512。（doi: 10.1126/science.286.5439.509）[8] Santos F C、Pacheco J M。2005， 无标度网络为合作的涌现提供统一框架。《物理评论快报》95，098104。（doi: 10.1103/PhysRevLett.95.098104）[9]Szab´o G、F´ath G。2007年，图上的演化博弈。《物理学报告》446，97-216。（doi: 10.1016/j.physrep.2007.04.004）[10] Tarnita C E、Ohtsuki H、Antal T、Fu F、Nowak M A。2009年，结构化种群中的策略选择。《理论生物学杂志》259，570-581。（doi: 10.1016/j.jtbi.2009.03.035）[11] Allen B、Lippner G、Chen Y T、Fotouhi B、Momeni N、Yau S T、Nowak M A。2017年，任意种群结构上的演化动力学。《自然》544，227-230。（doi: 10.1038/nature21723）[12] Li A、Zhou L、Su Q、Cornelius S P、Liu Y Y、Wang L、Levin S A。2020，时序网络上的合作演化。《自然·通讯》11，2259。（doi: 10.1038/s41467-020-16088-w）[13] Su Q、McAvoy A、Mori Y、Plotkin J B。2022，多层种群中亲社会行为的演化。《自然·人类行为》6，338-348。（doi: 10.1038/s41562-021-01241-2）[14] Su Q、Allen B、Plotkin J B。2022，与非对称社会互动相伴的合作演化。《美国国家科学院院刊》119，e2113468118。（doi: 10.1073/pnas.2113468118）

[15] Ohtsuki H、Hauert C、Lieberman E、Nowak M A。2006年。图和社交网络上合作演化的一个简单规则。《自然》441，502-505。（doi: 10.1038/nature04605）[16] Ohtsuki H、Nowak M A。2006年。图上的复制子方程。《理论生物学杂志》243，86-97。（doi: 10.1016/j.jtbi.2006.06.004）[17] Nowak M A、Tarnita C E、Antal T。2010年。结构化种群中的演化动力学。《皇家学会哲学汇刊B辑》365，19-30。（doi: 10.1098/rstb.2009.0215）[18] Zhou L、Wu B、Du J、Wang L。2021年。志向动力学在异质种群中产生稳健预测。《自然·通讯》12，3250。（doi: 10.1038/s41467-021-23548-4）[19] Capraro V、Perc M。2021年。道德偏好的数学基础。《皇家学会界面》18，20200880。（doi: 10.1098/rsif.2020.0880）[20] Capraro V、Halpern J Y、Perc M。2022年。从基于结果的偏好到基于语言的偏好。《经济文献杂志》。（即将出版）[21] Henrich J。2006年。合作、惩罚与人类制度的演化。《科学》312，60-61（2006）。（doi: 10.1126/science.1126398）.[22] G¨urerk O、Irlenbusch B、Rockenbach B。2006年。制裁制度的竞争优势。《科学》312，108-111。（doi: 10.1126/science.1123633）[23] Dreber A、Rand D G、Fudenberg D、Nowak M A。2008年。获胜者不进行惩罚。《自然》452，348-351。（doi: 10.1038/nature06723）[24] Rand D G、Dreber A、Ellingsen T、Fudenberg D、Nowak M A。2009年。积极互动促进公共合作。《科学》325，1272-1275。（doi: 10.1126/science.1177418）[25] Sigmund K、De Silva H、Traulsen A、Hauert C。2010年。社会学习促进公地治理制度的形成。《自然》466，861-863。（doi: 10.1038/nature09203）[26] Han T A、Pereira L M、Lenaerts T。2015年。在公共物品博弈中规避或限制背叛者？《皇家学会界面》12，20141203。（doi: 10.1098/rsif.2014.1203）[27] Mann R P、Helbing D。2017年。集体智慧的最优激励。《美国国家科学院院刊》114，5077-5082。（doi: 10.1073/pnas.1618722114）[28] Riehl J、Ramazi P、Cao M。2018年。二元决策网络中异步最佳响应动力学的基于激励的控制。《IEEE控制网络系统汇刊》6，727-736。（doi: 10.1109/TCNS.2018.2873166）[29] Vasconcelos V V、Dannenberg A、Levin S A。2022年。通过投票和学习选择并维持惩罚制度。《自然·可持续性》5，578-585。（doi: 10.1038/s41893-022-00877-w）

[30] Sasaki T、Br¨annstr¨om ˚A、Dieckmann U、Sigmund K。2012年。接受或拒绝选项使小额惩罚能够克服社会困境。《美国国家科学院院刊》109，1165-1169。（doi: 10.1073/pnas.1115219109）[31] Vasconcelos V V、Santos F C、Pacheco J M。2013年。对高风险公地进行合作治理的自下而上制度方法。《自然·气候变化》3，797-801。（doi: 10.1038/NCLIMATE1927）[32] Chen X、Sasaki T、Br¨annstr¨om ˚A、Dieckmann U。2015年。先胡萝卜，后大棒：激励措施的适应性混合如何促进合作。《皇家学会界面》12，20140935。（doi: 10.1098/rsif.2014.0935）[33] Wang S、Chen X、Szolnoki A。2019年。探索公共合作的最优制度激励。《非线性科学与数值模拟通讯》79，104914。（doi: 10.1016/j.cnsns.2019.104914）[34] Duong M H、Han T A。2021年。有限种群中促进合作的制度激励成本效率。《英国皇家学会会刊A辑》477，20210568。（doi: 10.1098/rspa.2021.0568）[35] Szab´o G、T˝oke C。1998年。方格晶格上的演化囚徒困境博弈。《物理评论E》58，69。（doi: 10.1103/PhysRevE.58.69）[36] Nowak M A、Sasaki A、Taylor C、Fudenberg D。2004年。有限种群中合作的出现与演化稳定性。《自然》428，646-650。（doi: 10.1038/nature02414）[37] Evans L C。2005年。数学最优控制理论导论。美国伯克利：加利福尼亚大学出版社。[38] Geering H P。2007年。具有工程应用的最优控制。德国柏林：施普林格。[39] Lenhart S、Workman J T。2007年。应用于生物模型的最优控制。美国博卡拉顿：Chapman and Hall/CRC。[40] Ostrom E。1990年。公地治理：集体行动制度的演化。美国纽约州纽约：剑桥大学出版社。[41] G¨achter S。2012年。胡萝卜还是大棒? 《自然》483，39-40。[42] Morita S。2008年。复杂网络上演化博弈的扩展成对近似。《理论物理进展》19，29-38。（doi: 10.1143/PTP.119.29）[43] Overton C E、Broom M、Hadjichrysanthou C、Sharkey K J。2019年。近似图上的随机演化动力学的方法。《理论生物学杂志》468，45-59。（doi: 10.1016/j.jtbi.2019.02.009）[44] Pinheiro F L、Santos F C、Pacheco J M。2012年。选择压力如何改变结构化种群中社会困境的性质。《新物理学杂志》14，073035。（doi: 10.1088/1367-2630/14/7/073035）[45] Zisis I、Di Guida S、Han T A、Kirchsteiger G、Lenaerts T。2015年。由接受驱动的慷慨-

预期博弈的演化分析。《科学报告》**5**，18076。（doi: 10.1038/srep18076）[46] McAvoy A、Rao A、Hauert C。2021年 选择强度对亲社会行为演化的有趣影响。《公共科学图书馆·计算生物学》**17**，e1009611。（doi: 10.1371/journal.pcbi.1009611）[47] Han T A、Tran-Thanh L。2018 促进合作演化的高成本效益外部干预。《科学报告》**8**，15997。（doi: 10.1038/s41598-018-34435-2）[48]Ibsen-Jensen R、Chatteriee K、Nowak M A。2015 生态与演化空间动力学的计算复杂性。《美国国家科学院院刊》 **112**，15636-15641。（doi: 10.1073/pnas.1511366112）[49] Doebeli M、Hauert C。2005 基于囚徒困境和雪堆博弈的合作模型。《生态学快报》**8**，748-766。（doi: 10.1111/j.1461-0248.2005.00773.x）[50] Skyrms B。2004 鹿猎博弈与社会结构的演化。 英国剑桥：剑桥大学出版社。[51] Perc M、G´omez-Garde˜nes J、Szolnoki A、Flor´ıa M、Moreno Y。2013 结构化种群中群体互动的演化动力学：综述。《皇家学会界面》 **10**，20120997。（doi: 10.1098/rsif.2012.0997）[52] Li A、Wu B、Wang L。2014 协同互动与局部互动并存时，合作可能不如单独存在任一互动时。《科学报告》**4**，5536。（doi: 10.1038/srep05536）[53] Li A、Broom M、Du J、Wang L。2016 结构化种群中一般群体互动的演化动力学。《物理评论*E*》**93**，022407。（doi: 10.1103/PhysRevE.93.022407）[54] Grilli J、Barab´as G、Michalska-Smith M J、Allesina S。2017 高阶互动稳定竞争网络模型中的动力学。《自然》**548**，210-213。（doi: 10.1038/nature23273）[55] Alvarez-Rodriguez U、Battiston F、Ferraz de Arruda G、Moreno Y、Perc M、Latora V。2021 社交网络中高阶互动的演化动力学。《自然·人类行为》**5**，586-595。（doi: 10.1038/s41562-020-01024-1）[56] Han T A、Lynch S、Tran-Thanh L、Santos F C。2018 通过局部和全局干预策略促进结构化种群中的合作。载于：第27届国际人工智能联合会议暨第23届欧洲人工智能会议论文集，289-295。[57] Cimpeanu T、Perret C、Han T A。2021 促进最后通牒博弈中公平的高成本效益干预。《知识工程系统》**233**，107545。（doi: 10.1016/j.knosys.2021.107545）[58] Cimpeanu T、Han T A、Santos F C。2019 通过外生奖励促进无标度网络中的合作。载于：2019年人工生命会议论文集，316-323。

[59] Capraro V、Perc M、Vilone D。2020 网络中的说谎：结构和拓扑在促进诚实方面的作用。《物理评论*E*》 **101**，032305。（doi: 10.1103/PhysRevE.101.032305）

[60] Kumar A、Capraro V、Perc M。2020《信任与可信赖性的演化》。《皇家学会界面》**17**, 20200491。（doi: 10.1098/rsif.2020.0491）

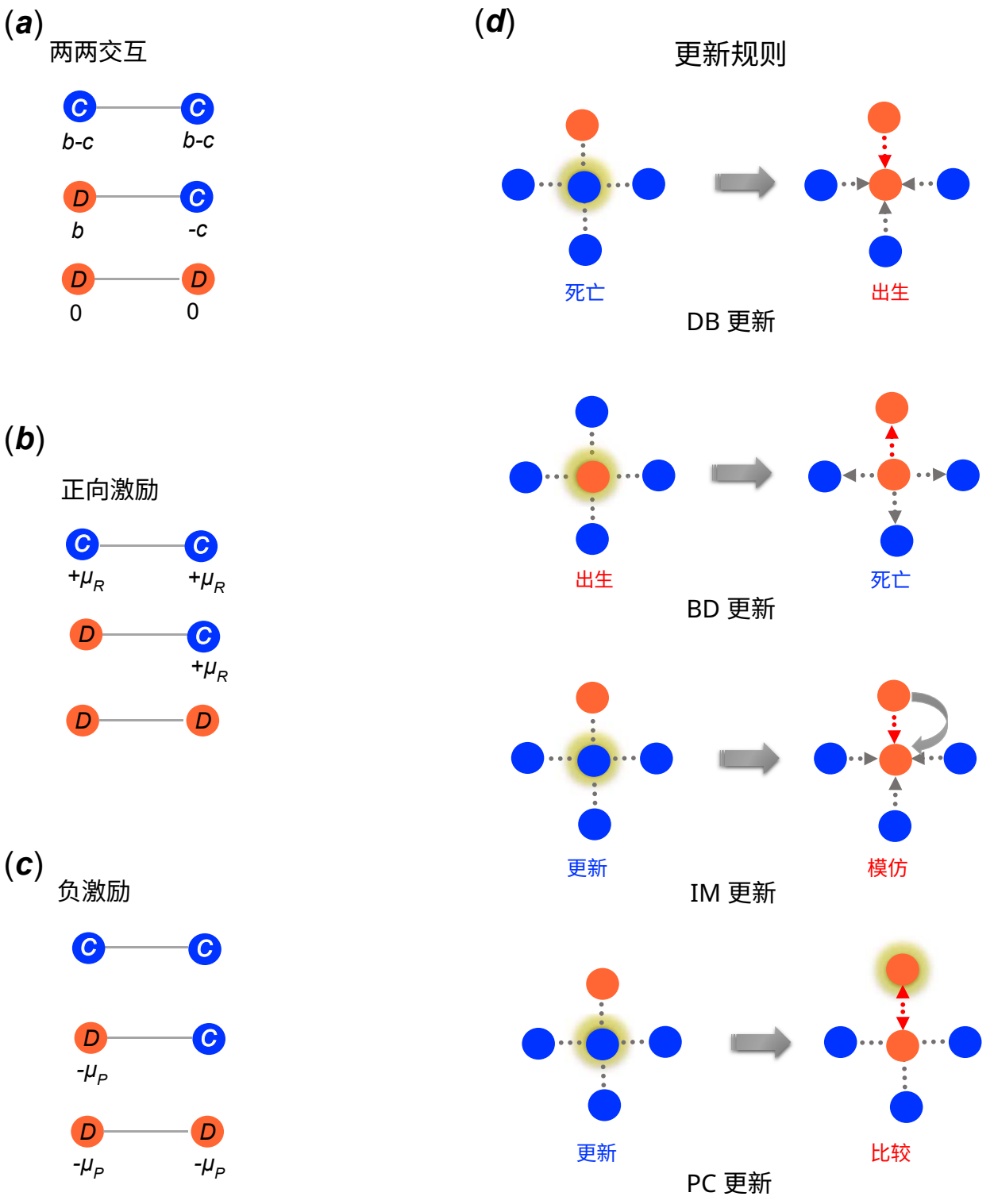

图 1：具有制度奖励或惩罚的图上演化囚徒困境博弈。

面板 a 展示了网络中两个相连邻居之间的两两交互。面板 b（c）展示了对参与博弈的两个相连智能体实施激励的方式，分别对应激励提供机构施加正向（负向）激励的情形。面板 d 展示了四种策略更新规则，说明个体在从与邻居及激励提供机构的两两交互中获得收益后，如何更新策略。

更新规则	最小 μ_R / μ_P	μ_R^* / μ_P^*	J_R^*	J_P^*
DB 更新	$c-\frac{b}{k}$	$\frac{2(kc-b)}{k}$	$\frac{(kN\mu_R^*)^2}{2\beta_{DB}}[p_0+\delta-1+\ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{DB}}[p_0+\delta-1+\ln(\frac{1-\delta}{p_0})]$
BD 更新	c	$2c$	$\frac{(kN\mu_R^*)^2}{2\beta_{BD}}[p_0+\delta-1+\ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{BD}}[p_0+\delta-1+\ln(\frac{1-\delta}{p_0})]$
IM 更新	$c-\frac{b}{k+2}$	$\frac{2(k+2)c-2b}{k+2}$	$\frac{(kN\mu_R^*)^2}{2\beta_{IM}}[p_0+\delta-1+\ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{IM}}[p_0+\delta-1+\ln(\frac{1-\delta}{p_0})]$
PC 更新	c	$2c$	$\frac{(kN\mu_R^*)^2}{2\beta_{PC}}[p_0+\delta-1+\ln(\frac{1-p_0}{\delta})]$	$\frac{(kN\mu_P^*)^2}{2\beta_{PC}}[p_0+\delta-1+\ln(\frac{1-\delta}{p_0})]$

图2：不同策略更新规则下合作的制度激励优化。

这里，最小的 μ_R （μ_P）分别表示在不同策略更新规则下，为促进合作演化所需的最小正向（负向）激励量。 μ_R^* （μ_P^*）分别表示不同策略更新规则下的最优正向（负向）激励方案。 J_R^*（J_P^*）分别表示采用最优奖励（惩罚）方案 μ_R^* （μ_P^*）后，动力系统从初始状态 p_0到达预期终止状态 $1-\delta$ 所产生的累积成本。此外， N 表示种群规模， k 表示正则网络的度。 b 表示合作收益， c 表示合作成本。在 DB 更新下的参数为$\beta_{DB}=\frac{\omega}{(k-2)(ck-b)^{k-1}}$，在 BD 更新下为 $\beta_{BD}=\frac{\omega k}{(k-2)^{k-1}c}$，在 IM 更新下为 $\beta_{IM}=\frac{\omega k^2}{(k-2)[c(k+2)-b](k+1)^2(k-1)}$，在 PC 更新下为 $\beta_{PC}=\frac{\omega k}{(k-2)2^{k-1}c}$（）。

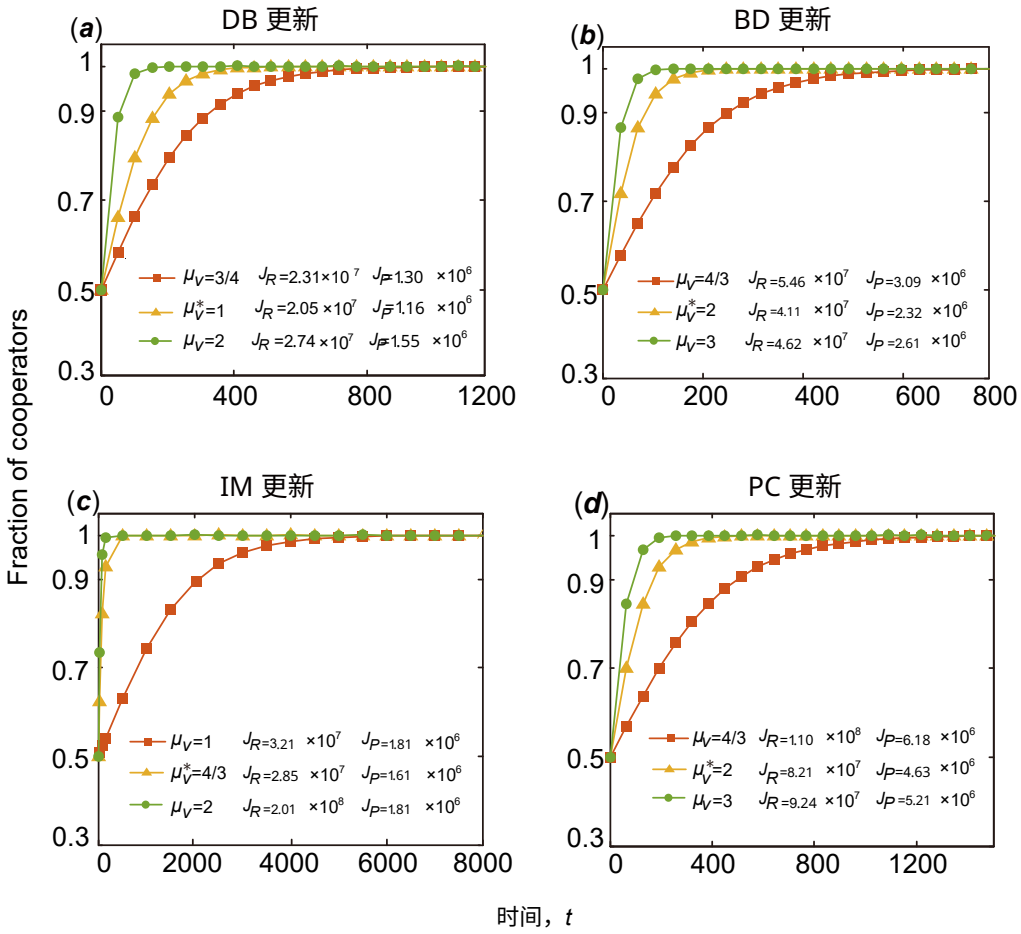

图 3：不同策略更新规则下，正向和负向激励对应的合作者比例随时间的演化。 每个面板展示了基于所得动力学方程，通过数值计算得出的结果，涵盖奖励（R）或惩罚（P）的不同激励水平。最优激励水平以 $*$ 标出。为便于比较，我们还标出了每种激励方案达到期望终止状态所需的 J_R 和 J_P 累积成本。参数： $N = 100$、 $b = 2$、 $c = 1$、$\delta = 0.01$、 $\omega = 0.01$、 $p_0 = 0.5$和 $k = 4$。

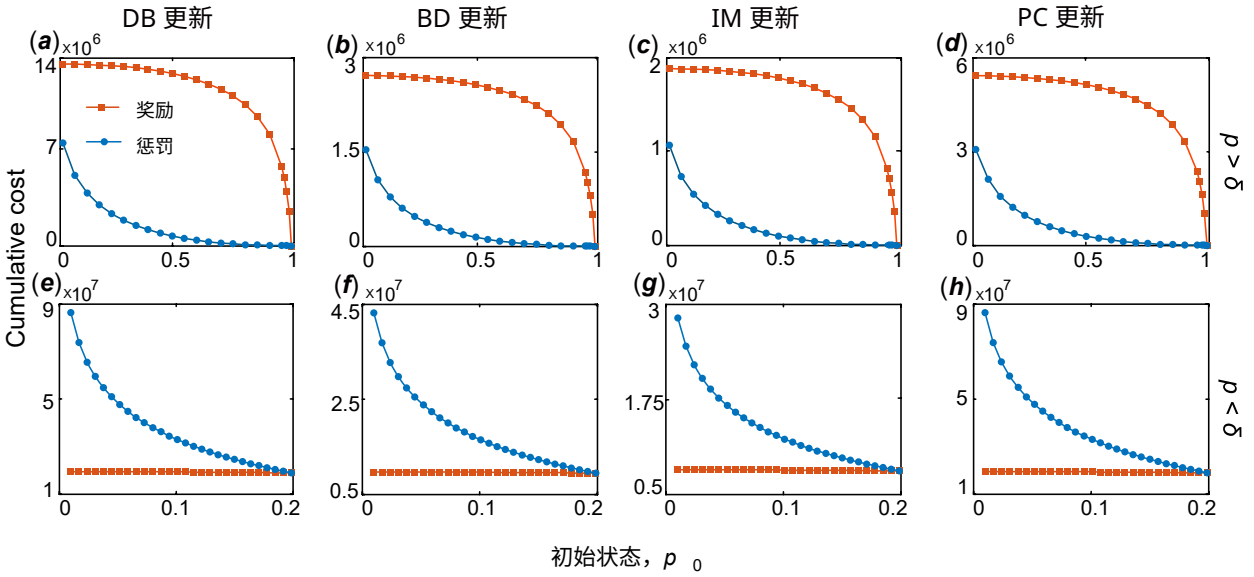

图 4：达到预期终止状态所需的累积成本与 p_0最优奖励和惩罚协议中合作者的初始比例以及不同策略更新规则的关系。每列面板代表一种策略更新规则，具体如图所示。顶行表示在条件 0.01 下，基于所得动力学方程进行数值计算的结果；底行表示在条件 $p_0 < \delta = 0.2$ 下进行数值计算的结果。其他参数： $N = 100$、 $b = 2$、 $c = 1$、 $\omega = 0.01$ 和 $k = 4$。

电子补充材料

优　　化结构化种群中促进合作的制度激励

Shengxian Wang、Xiaojie Chen、Zhilong Xiao、Attila Szolnoki 和 Vı́tor V. Vasconcelos

在电子补充材料中，我们提供了详细的理论分析，以探讨促进结构化种群中合作的最优激励方案。具体而言，我们考虑了四种不同的策略更新规则：第 1 节介绍数据库更新，第 2 节介绍 BD 更新，第 3 节介绍 IM 更新，最后第 4 节介绍 PC 更新。在每一节中，我们首先使用成对近似法，探讨存在正激励或负激励时的动力学方程，并从理论上得出合作演化所需的最低激励量条件。随后，我们分别针对正、负激励方案构建最优控制问题，并通过 HJB方程的方法求得最优的正、负激励方案。因此，达到预期最终状态所需的最优激励方案累积成本也得以确定。

1. 数据库更新

1.1. 正向激励

种群结构由一个度为 $k > 2$ 的 N 节点正则网络表示。网络的顶点对应个体，边表示个体之间的互动关系。每个个体都与其邻居进行囚徒困境博弈，策略可以是合作（C）或背叛（D）。下面介绍一些符号。令 p_i 表示采用策略 i 的个体比例，令 p_{ij} 表示 ij–对的比例；最后，令 $q_{i|j}$ 表示在相邻节点为 j–个体的条件下，找到一个 i–个体的条件概率，其中$i,\ j \in \{C, D\}$。利用这些符号，我们有 $p_C + p_D = 1$、 $q_{C|i} + q_{D|i} = 1$、 $p_{ij} = q_{i|j}p_j$ 和 $p_{CD} = p_{DC}$。

我们首先考虑将正向激励引入采用 DB 更新的网络化囚徒困境博弈 [15, 16]。根据 DB 更新，我们以概率 p_i ($i \in \{C, D\}$) 随机选择一个焦点个体使其死亡，其中 i 表示该焦点个体的策略。令 k_C 和 k_D分别表示其 k 个邻居中合作者和背叛者的数量，且 $k_D + k_C = k$。如果该

焦点个体采用策略 D，那么一个 C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)(k-1)q_{C|C} + (\mu_R - c)[(k-1)q_{D|C} + 1]\}, \tag{S1}$$

且 a D–邻居的适应度为

$$f_D = 1 - \omega + \omega[b(k-1)q_{C|D}], \tag{S2}$$

其中， $0 \leq \omega \leq 1$衡量选择强度。

由于焦点个体的所有邻居以与其适应度成正比的概率争夺空位，因此 a C–邻居取代该空位的概率为

$$\Gamma = \frac{k_C f_C}{k_C f_C + k_D f_D}. \tag{S3}$$

根据上述方程， p_C 以概率 $1/N$ 增加

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma, \tag{S4}$$

且 CC-对的数量以 k_C 增加，因此 p_{CC} 以概率 $k_C/(kN/2)$增加

y

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma. \tag{S5}$$

此外，我们考虑另一种情况，即随机选中的焦点个体采取策略 C。在这种情况下，a C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)[(k-1)q_{C|C} + 1] + (\mu_R - c)(k-1)q_{D|C}\}, \tag{S6}$$

而 a D–邻居的适应度为

$$f_D = 1 - \omega + \omega\{b[(k-1)q_{C|D} + 1] + 0 \cdot (k-1)q_{D|D}\}. \tag{S7}$$

a D–邻居取代空位的概率为

$$M = \frac{k_D f_D}{k_C f_C + k_D f_D}. \tag{S8}$$

因此， p_C 以概率降低 $1/N$

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M, \tag{S9}$$

并且 CC–对的数量减少 k_C，因此 p_{CC} 以概率降低 $k_C/（kN/2）$

能力

$$P(\Delta p_{CC}=-\frac{2k_C}{kN})=p_C\binom{k}{k_C}(q_{C|D})^{k_C}(q_{D|D})^{k_D}M. \tag{S10}$$

我们假设每个时间单位发生一次替换事件，且 p_C 的导数可写为

$$\begin{aligned}\frac{dp_C}{dt}&=\frac{E(\Delta p_C)}{\Delta t}=\frac{\frac{1}{N}P(\Delta p_C=\frac{1}{N})-\frac{1}{N}P(\Delta p_C=-\frac{1}{N})}{\frac{1}{N}}\\&=\omega\frac{k-1}{k}p_{CD}[(\mu_R-c)+(k-1)\eta_1](q_{C|C}+q_{D|D})+o(\omega^2),\end{aligned} \tag{S11}$$

其中，$\eta_1=(b-c+\mu_R)q_{C|C}+(\mu_R-c)q_{D|C}-bq_{C|D}$。相应地，$p_{CC}$ 的导数为

$$\begin{aligned}\frac{dp_{CC}}{dt}&=\frac{E(\Delta p_{CC})}{\Delta t}=\frac{\sum_{k_C=0}^{k}\frac{2k_C}{kN}P(\Delta p_{CC}=\frac{2k_C}{kN})-\sum_{k_C=0}^{k}\frac{2k_C}{kN}P(\Delta p_{CC}=-\frac{2k_C}{kN})}{\frac{1}{N}}\\&=\frac{2p_{CD}}{k}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega).\end{aligned} \tag{S12}$$

由于 $q_{C|C}=\frac{p_{CC}}{p_C}$，我们有

$$\frac{dq_{C|C}}{dt}=\frac{d}{dt}(\frac{p_{CC}}{p_C})=\frac{2p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega). \tag{S13}$$

其他变量，例如 p_D 和$q_{D|C}$，也可以通过适当计算表示为 p_C 和$q_{C|C}$ 的函数，随后动力系统可由 p_C 和$q_{C|C}$描述。将方程（S11）和（S13）的右侧表达式改写为 p_C 和$q_{C|C}$ 的函数，可得动力学方程

$$\begin{cases}\frac{dp_C}{dt}=\omega\Psi_{\mathrm{DB}}^{R}(p_C,q_{C|C})+o(\omega^2),\\\frac{dq_{C|C}}{dt}=\Phi_{\mathrm{DB}}^{R}(p_C,q_{C|C})+o(\omega),\end{cases} \tag{S14}$$

其中

$$\begin{cases}\Psi_{\mathrm{DB}}^{R}(p_C,q_{C|C})=\frac{k-1}{k}p_{CD}[(-c+\mu_R)+(k-1)\eta_1](q_{C|C}+q_{D|D}),\\\Phi_{\mathrm{DB}}^{R}(p_C,q_{C|C})=\frac{2p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})].\end{cases} \tag{S15}$$

在弱选择（$w\ll 1$）下，$q_{C|C}$ 的速度可能很快，并且可能迅速收敛到由 $\Phi_{\mathrm{DB}}^{R}(p_C,q_{C|C})$ 定义的根 $=0$，随着时间 $t\to+\infty$。因此，我们得到

$$q_{C|C}=p_C+\frac{1}{k-1}(1-p_C). \tag{S16}$$

因此，方程（S14）所描述的动力学方程变为

$$\frac{dp_C}{dt}=\frac{\omega(k-2)[b+k(\mu_R-c)]}{k-1}p_C(1-p_C)+o(\omega^2), \tag{S17}$$

其具有 t 两个不动点 $p_C = 0$ 和 $p_C = 1$。我们将函数 $F_{\mathrm{DB}}(p_C, \mu_R, t)$ 定义为

$$F_{\mathrm{DB}}(p_C, \mu_R, t) = \frac{\omega(k-2)[b + k(\mu_R - c)]}{k-1} p_C(1-p_C) + o(\omega^2). \tag{S18}$$

该函数是连续可微函数， $F_{\mathrm{DB}}(p_C, \mu_R, t)$ 对 p_C 的导数为

$$\frac{dF_{\mathrm{DB}}}{dp_C} = \frac{\omega(k-2)[b + k(\mu_R - c)]}{k-1}(1-2p_C) + o(\omega^2). \tag{S19}$$

对于 $\mu_R > c - \frac{b}{k}$，我们有 $\frac{dF_{\mathrm{DB}}}{dp_C}|_{p_C=1} = -\frac{\omega(k-2)[b+k(\mu_R-c)}{)_{k-1}} \quad \frac{]}{} < 0$ 且 $\frac{dF_{\mathrm{DB}}}{dp_C}|_{p_C=0} = \frac{\omega}{(k-2)}$ $[b+k\ \mu_R - c\ ]\ >()0$。这意味着不动点 $p_C = 1$ 是稳定的，而 $p_C = 0$ 不稳定，即合作者胜过背叛者。

最后，我们研究 $\mu_R = 0$这一特殊情况。在这种情况下，可以看出，对于 $b/c > k$，不动点 $p_C = 1$ 是稳定的，而 $p_C = 0$ 不稳定。因此，我们得到了合作演化的条件 $b/c > k$，与参考文献 [15, 16]中先前得到的结果一致。

1.2. 负激励

在本小节中，我们接着考虑将负激励引入采用 DB 更新的网络化囚徒困境博弈，收益矩阵由正文中的方程 (15) 给出。根据 DB 更新，如果焦点个体采用策略 D，那么 C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b-c)(k-1)q_{C|C} - c[(k-1)q_{D|C} + 1]\}, \tag{S20}$$

而 D–邻居的适应度为

$$f_D = 1 - \omega + \omega\{(b-\mu_P)(k-1)q_{C|D} - \mu_P[(k-1)q_{D|D} + 1]\}. \tag{S21}$$

C–邻居取代空位的概率由方程 (S3) 中的表达式 Γ 给出。因此， p_C 以概率 $1/N$ 增加

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma. \tag{S22}$$

相应地， CC–对的数量增加 k_C，因此 p_{CC} 以概率 $k_C/（kN/2）$ 增加

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Gamma. \tag{S23}$$

此外，我们还考虑焦点个体采取策略 D的另一种情况。在这种情况下， C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b-c)[(k-1)q_{C|C} + 1] - c(k-1)q_{D|C}\}, \tag{S24}$$

而 D–邻居的适应度为

$$f_D = 1 - \omega + \omega\{(b - \mu_P)[(k-1)q_{C|D} + 1] - \mu_P(k-1)q_{D|D}\}. \tag{S25}$$

D–邻居取代空位的概率也可以用方程（S8）中的表达式 M给出。因此， p_C 以概率 $1/N$ 减少

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M. \tag{S26}$$

相应地， CC–对的数量减少 k_C ，且 p_{CC} 以概率 $k_C/(kN/2)$ 减少

$$P(\Delta p_{CC} = -\frac{2k_C}{kN}) = p_C \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} M. \tag{S27}$$

根据这些计算，我们得到 p_C 的时间导数为

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega(k-1)}{k} p_{CD}[\mu_P - c + (k-1)\eta_2](q_{C|C} + q_{D|D}) + o(\omega^2), \end{aligned} \tag{S28}$$

其中 $\eta_2 = (b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}$。 而 p_{CC} 的时间导数为

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = \frac{2k_C}{kN}) - \sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = -\frac{2k_C}{kN})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S29}$$

此外，我们有

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \tag{S30}$$

因此，动力学方程可以表示为

$$\begin{cases} \frac{dp_C}{dt} = \omega \Psi^P_{\mathrm{DB}}(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi^P_{\mathrm{DB}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S31}$$

其中

$$\begin{cases} \Psi^P_{\mathrm{DB}}(p_C, q_{C|C}) = \frac{k-1}{k} p_{CD}[\mu_P - c + (k-1)\eta_2](q_{C|C} + q_{D|D}), \\ \Phi^P_{\mathrm{DB}}(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很快，并且可能迅速收敛到由 $\Phi^P_{\mathrm{DB}}(p_C, q_{C|C}) = 0$ 定义的根，随着时间 $t \to +\infty$。因此，我们得到

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S32}$$

相应地，方程 (S31) 所描述的动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega(k-2)[b + k(\mu_P - c)]}{k-1} p_C(1 - p_C) + o(\omega^2), \tag{S33}$$

它有两个不动点 $p_C = 0$ 和 $p_C = 1$。我们将函数 $F_{\mathrm{DB}}(p_C, \mu_P, t)$ 定义为

$$F_{\mathrm{DB}}(p_C, \mu_P, t) = \frac{\omega(k-2)[b + k(\mu_P - c)]}{k-1} p_C(1 - p_C) + o(\omega^2), \tag{S34}$$

$F_{\mathrm{DB}}(p_C, \mu_P, t)$ 对 p_C 的导数为

$$\frac{dF_{\mathrm{DB}}}{dp_C} = \frac{\omega(k-2)[b + k(\mu_P - c)]}{k-1}(1 - 2p_C) + o(\omega^2). \tag{S35}$$

因此，对于 $\mu_P > c - \frac{b}{k}$，有 $\frac{dF_{\mathrm{DB}}}{dp_C}\big|_{p_C=1} = -\frac{\omega(k-2)[b+k(\mu_P - c)}{k-1}] < 0$ 且 $\frac{dF_{\mathrm{DB}}}{dp_C}\big|_{p_C=0} = \frac{\omega(k-2)[b+k\ \mu_P - c\ ]}{k-1} > 0$()，这意味着不动点 $p_C = 1$ 稳定而 $p_C = 0$ 不稳定，即合作者胜过背叛者。特别地，当 $\mu_P = 0$ 时，可以看出对于 $b/c > k$，不动点 $p_C = 1$ 稳定而 $p_C = 0$ 不稳定。因此，我们得到文献 [15, 16] 中关于合作演化的条件 $b/c > k$。

1.3. 最优激励协议

在第 1.1 和 1.2 小节中，我们通过弱选择极限下的成对近似方法，从理论上推导出了具有正激励或负激励的动力系统，其表达式为

$$\frac{dp_C}{dt} = F_{\mathrm{DB}}(p_C, \mu_v, t) = \frac{\omega(k-2)[b + k(\mu_v - c)]}{k-1} p_C(1 - p_C) + o(\omega^2). \tag{S36}$$

如前所述，该动力系统有两个均衡点，分别为 $p_C = 0$ 和 $p_C = 1$。如果 $\mu_v > c - \frac{b}{k}$，前者不稳定，后者稳定，这意味着从长远来看，合作将得到促进。由于提供激励需要付出成本，我们的主要目标是探究能够促进合作且成本最低的最优激励方案。为此，我们现在求解针对 DB 更新所建立的最优控制问题。

首先，我们求解奖励的最优控制问题。哈密顿函数$H_{\mathrm{DB}}(p_C, \mu_R, t)$定义为

$$H_{\mathrm{DB}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J^*_R}{\partial p_C} F_{\mathrm{DB}}(p_C, \mu_R, t), \tag{S37}$$

其中， J_R^* 是针对最优奖励方案的 p_C 和 t 的最优成本函数，定义为

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R^*)^2}{2} dt. \tag{S38}$$

通过求解 $\frac{\partial H_{\mathrm{DB}}}{\partial \mu_R} = 0$，我们知道最优奖励方案$\mu_R^*$ 应满足

$$\mu_R^* = -\frac{\omega(k-2)(1-p_C)}{N^2k(k-1)p_C} \frac{\partial J_R^*}{\partial p_C}. \tag{S39}$$

通常，我们应求解方程 (S37) 的正则方程，以得到最优奖励方案 [37–39]。然而，所得动力系统是非线性的，这大大增加了通过直接计算获得最优方案精确表达式的复杂性。相反，为求解最优控制问题，我们采用连续时间系统的动态规划方法——HJB方程 [37–39]。该方程可写为

$$-\frac{\partial J_R^*}{\partial t} = H_{\mathrm{DB}}(p_C, \mu_R^*, t). \tag{S40}$$

将方程 (S39) 代入上述 HJB方程，我们得到

$$-\frac{\partial J_R^*}{\partial t} = \frac{(kNp_C\mu_R^*)^2}{2} + \frac{\omega(k-2)[b+k(\mu_R^*-c)]}{k-1} p_C(1-p_C)\frac{\partial J_R^*}{\partial p_C}. \tag{S41}$$

由于我们假设终止时间 t_f 不固定，最优成本函数 $J_R^*(p_C,\ t)$ 与 t无关。因此，我们有

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S42}$$

于是得到

$$\frac{\partial J_R^*}{\partial p_C} = 0 \ \text{ or } \ \frac{\partial J_R^*}{\partial p_C} = \frac{2N^2(k-1)(b-ck)p_C}{\omega(k-2)(1-p_C)}. \tag{S43}$$

由于 $\mu_R > 0$ 和 $p_C \in (0,\ 1)$，我们有

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S44}$$

由方程 (S44) 可知，对于 $b/c \geq k$，该不等式显然成立。因此，我们考虑 $b/c < k$ 这一情形，此时只有 $\frac{\partial J_R^*}{\partial p_C} = {}^{2N^2}{}_{(k-1)(b-ck)\omega\ k-2}$ $\underline{{}^{p_C}{}_{()(1-p_C)}}$ 成立。将该方程代入方程 (S39)，可得最优奖励水平为

$$\mu_R^* = \frac{2(ck-b)}{k}. \tag{S45}$$

据此 μ_R^* 动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega(k-2)(ck-b)}{k-1} p_C(1-p_C), \tag{S46}$$

其中，种群中合作者初始比例记为 $p_0 = p_C(0)$。该方程的解为

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\mathrm{DB}} t}}, \tag{S47}$$

其中 $\beta_{\mathrm{DB}} = \frac{\omega}{(k-2)(ck-b)\,k-1}$。这也意味着，动力系统需要无限长的时间才能从初始 $p_0 < 1$ 到达完全合作状态。为避免这种情况，我们假设终止状态 $p_C(t_f)$ 为 $1-\delta$，其中 δ是决定终止时间合作水平的参数。由于 $b/c < k$， p_C会随时间 t单调递增，从而导致 $p_C(t_f) > p_0$。

此外，动力系统到达预期终止状态 $p_C(t_f)$ 所需的最优奖励水平 μ_R^* 对应的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\mathrm{DB}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S48}$$

惩罚的最优控制问题也可以用类似的方法求解，对于 μ_P^* 最优水平，我们有

$$\mu_P^* = \frac{2}{k}(ck - b) \tag{S49}$$

and

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\mathrm{DB}} t}}. \tag{S50}$$

因此，最优惩罚方案 μ_P^* 产生的累积成本为

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\mathrm{DB}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S51}$$

由这些结果可知，负向激励和正向激励的最优水平相同，即 $\mu_R^* = \mu_P^*$，但它们的累积成本可能不同。为了进行适当的比较，我们可以计算二者之差，即

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\mathrm{DB}}} \ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S52}$$

其中 $\mu_v^* = \mu_R^* = \mu_P^*$。在本研究中，我们假设 $p_0 > 0$，因为我们不考虑行为突变或策略更新错误。此外， $p_0 < p_C$（t_f） $= 1-\delta$，并且我们有 $p_0 + \delta < 1$。因此，当 $\delta < p_0$ 时，我们有 $J_R^* > J_P^*$，这意味着对于 DB 更新，最优惩罚所需的累积成本始终低于采用最优奖励所需的成本。但当 $\delta > p_0$ 时，我们有 $J_R^* < J_P^*$，这意味着对于 DB 更新，最优奖励所需的累积成本始终低于采用最优惩罚所需的成本。这一观察结果得到了数值计算和蒙特卡洛模拟的支持，分别如图 4 和图 S5 所示。

2. BD 更新

2.1. 正向激励

根据 BD 更新规则 [15, 16], , 我们随机选择一个焦点个体进行繁殖，选择概率与其适应度成正比。该个体的 k 个邻居中有 k_C 个合作者和 k_D 个背叛者。如果焦点个体采用策略 C，则其适应度为

$$f_C = 1 - \omega + \omega[(b-c)k_C - ck_D]. \tag{S53}$$

由于被选中个体的后代会随机取代其一个邻居，因此 p_C 增加 $1/N$ 的概率为

$$P(\Delta p_C = \frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}, \tag{S54}$$

其中， $\bar{f}$ 表示整个种群的平均适应度。在这种情况下， CC– 对的数量增加 $(k-1)q_{C|D}+1$，而 p_{CC} 以概率 $[(k-1)q_{C|D}+1]/(kN/2)$ 增加

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}. \tag{S55}$$

在另一种情况下，随机选中的焦点个体采用策略 D。此时，焦点个体的适应度为

$$f_D = 1 - \omega + \omega(bk_C + 0 \cdot k_D),$$

因此， p_C 以概率 $1/N$ 减少

$$P(\Delta p_C = -\frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S56}$$

因此， CC–对的数量以概率减少（$k-1$）$q_{C|C}$，而 p_{CC} 减少（$k-1$）$q_{C|C}/$（$kN/2$）

$$P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S57}$$

因此，这里将整个种群的平均适应度记为

$$\begin{aligned} \bar{f} &= p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} f_C + p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} f_D \\ &= 1 - \omega + k\omega[(b-c+\mu_R)p_{CC} + (\mu_R - c)p_{CD} + bp_{CD}]. \end{aligned} \tag{S58}$$

基于 t　通过这些计算，我们分别得到 p_C 和 p_{CC} 的时间导数:

$$\begin{aligned}\frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega p_{CD}}{\bar{f}}\{(\mu_R - c - b) + (k-1)[(b - c + \mu_R)q_{C|C} + (\mu_R - c)q_{D|C} - bq_{C|D}]\} + o(\omega^2)\end{aligned} \tag{S59}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) - \frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega).\end{aligned} \tag{S60}$$

此外，我们有

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega). \tag{S61}$$

因此，动力学方程可表示为

$$\begin{cases} \frac{dp_C}{dt} = \omega\Psi^R_{\text{BD}}(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi^R_{\text{BD}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S62}$$

其中

$$\begin{cases} \Psi^R_{\text{BD}}(p_C, q_{C|C}) = \frac{p_{CD}}{\bar{f}}\{(\mu_R - c - b) + (k-1)[(b - c + \mu_R)q_{C|C} + (\mu_R - c)q_{D|C} - bq_{C|D}]\}, \\ \Phi^R_{\text{BD}}(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1]. \end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很大，并且可能迅速收敛到由 $\Phi^R_{\text{BD}}(p_C, q_{C|C}) = 0$ 定义的根，随着时间 $t \to +\infty$。因此，我们得到

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S63}$$

相应地，由公式（S62）描述的动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)(\mu_R - c)}{k-1}p_C(1 - p_C) + o(\omega^2), \tag{S64}$$

它有两个不动点 $p_C = 0$ 和 $p_C = 1$。我们将函数 $F_{\text{BD}}(p_C, \mu_R, t)$ 定义为

$$F_{\text{BD}}(p_C, \mu_R, t) = \frac{\omega k(k-2)(\mu_R - c)}{k-1}p_C(1 - p_C) + o(\omega^2), \tag{S65}$$

以及 $F_{\text{BD}}(p_C, \mu_R, t)$ 对 p_C 的导数为

$$\frac{dF_{\text{BD}}}{dp_C} = \frac{\omega k(k-2)(\mu_R - c)}{k-1}(1 - 2p_C) + o(\omega^2). \tag{S66}$$

对于 $\mu_R > c$，我们有 $\frac{dF_{\mathrm{BD}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_R-c)}{k-1} < 0$ 和 $\frac{dF_{\mathrm{BD}}}{dp_C}|_{p_C=0} = \frac{\omega k}{k-1}(k-2)(\mu_R-c) > 0$。这意味着不动点 $p_C = 1$ 稳定，而 $p_C = 0$ 不稳定，即合作者战胜背叛者。特别地，当 $\mu_R = 0$ 时，我们可以看到不动点 $p_C = 0$ 总是稳定的，而 $p_C = 1$ 不稳定，这意味着正如参考文献 [15, 16] 所得出的，合作无法在 BD 更新规则下涌现。

2.2. 负激励

在本小节中，我们考虑将负激励引入采用 BD 更新的网络化囚徒困境博弈。根据 BD 更新，随机选择一个焦点个体进行繁殖，该个体的 k 个邻居中有 k_C 个合作者和 k_D 个背叛者。这里，我们首先考虑焦点个体采用策略 C 的情况。此时，焦点个体的适应度为

$$f_C = 1 - \omega + \omega[k_C(b-c) + k_D(-c)], \tag{S67}$$

因此，p_C 以概率 $1/N$ 增加

$$P(\Delta p_C = \frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}, \tag{S68}$$

其中，$\bar{f}$ 表示整个种群的平均适应度。CC–对的数量增加了 $(k-1)q_{C|D}+1$，因此 p_{CC} 以概率 $kN/2$ 增加 $[(k-1)q_{C|D}+1]/$

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{f_C}{\bar{f}} \frac{k_D}{k}. \tag{S69}$$

此外，我们考虑另一种情况，即随机选中的焦点个体采用策略 D。在这种情况下，焦点个体的适应度为

$$f_D = 1 - \omega + \omega[k_C(b-\mu_P) + k_D(-\mu_P)],$$

因此，p_C 以概率 $1/N$ 减少

$$P(\Delta p_C = -\frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S70}$$

并且，CC–对的数量以概率 $(k-1)q_{C|C}$ 减少，因此 p_{CC} 减少 $(k-1)q_{C|C}/(kN/2)$

$$P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{f_D}{\bar{f}} \frac{k_C}{k}. \tag{S71}$$

这里，整个种群的平均适应度可通过以下方式计算:

$$\begin{aligned}\bar{f} &= p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} f_C + p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} f_D \\ &= 1 - \omega + \omega k[(b-c)p_{CC} - cp_{CD} + (b-\mu_P)p_{CD} - \mu_P p_{DD}].\end{aligned} \tag{S72}$$

From 通过这些计算，我们分别得到 p_C 和 p_{CC} 的时间导数a s

$$\begin{aligned}\frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C = \frac{1}{N}) - \frac{1}{N}P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega p_{CD}}{\bar{f}}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\} + o(\omega^2),\end{aligned} \tag{S73}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) - \frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}} \\ &= \frac{2}{k}p_{CD}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega).\end{aligned} \tag{S74}$$

此外，我们有

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1] + o(\omega). \tag{S75}$$

因此，动力学方程可表示为

$$\begin{cases} \frac{dp_C}{dt} = \omega\Psi^P_{\text{BD}}(p_C, q_{C|C}) + o(\omega^2), \\ fracdq_{C|C}dt = \Phi^P_{\text{BD}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S76}$$

其中

$$\begin{cases} \Psi^P_{\text{BD}}(p_C, q_{C|C}) = \frac{p_{CD}}{\bar{f}}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\}, \\ \Phi^P_{\text{BD}}(p_C, q_{C|C}) = \frac{2p_{CD}}{kp_C}[(k-1)(q_{C|D} - q_{C|C}) + 1]. \end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很大，并且可能迅速收敛到由 $\Phi^P_{\text{BD}}(p_C, q_{C|C}) = 0$ 定义的根，因为时间 $t \to +\infty$。因此，我们得到

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S77}$$

因此，由方程 (S76) 描述的系统变为

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)(\mu_P - c)}{k-1}p_C(1 - p_C) + o(\omega^2), \tag{S78}$$

它有两个不动点， $p_C = 0$和 $p_C = 1$。我们将函数 $F_{\mathrm{BD}}(p_C, \mu_P, t)$ 定义为

$$F_{\mathrm{BD}}(p_C, \mu_P, t) = \frac{\omega k(k-2)(\mu_P - c)}{k-1} p_C(1-p_C) + o(\omega^2), \tag{S79}$$

以及 $F_{\mathrm{BD}}(p_C, \mu_P, t)$ 对 p_C的导数为

$$\frac{dF_{\mathrm{BD}}}{dp_C} = \frac{\omega k(k-2)(\mu_P - c)}{k-1}(1-2p_C) + o(\omega^2). \tag{S80}$$

对于 $\mu_P > c$，我们有$\frac{dF_{\mathrm{BD}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_P - c)}{k-1} < 0$和 $\frac{dF_{\mathrm{BD}}}{dp_C}|_{p_C=0} = \frac{\omega k}{(k-2)}(\mu_P - c)_{k-1} > 0$，这意味着不动点 $p_C = 1$ 是稳定的，而 $p_C = 0$ 是不稳定的，即合作者胜过背叛者。特别地，当 $\mu_P = 0$时，我们可以看到，不动点 $p_C = 0$ 始终稳定，而 $p_C = 1$ 不稳定，这意味着按照 BD更新规则，合作永远无法涌现，这与参考文献 [15, 16]中的结论一致。

2.3. 最优激励协议

在第 2.1 和 2.2 小节中，我们利用弱选择极限下的成对近似方法，从理论上得到了存在正激励或负激励时的动力学方程，其表达式为

$$\frac{dp_C}{dt} = F_{\mathrm{BD}}(p_C, \mu_v, t) = \frac{\omega k(k-2)(\mu_v - c)}{k-1} p_C(1-p_C) + o(\omega^2)\,, \tag{S81}$$

具有 $p_C = 0$ 和 $p_C = 1$ 两个不动点。若 $\mu_v > c$，前者不稳定，后者稳定，这表明从长远来看，合作将得到促进。此外，为了探究最优奖励与惩罚方案，我们现在采用 HJB方程的方法，求解针对这种 BD 更新所建立的最优控制问题。

首先，我们求解奖励的最优控制问题。我们将哈密顿函数$H_{\mathrm{BD}}(p_C, \mu_R, t)$定义为

$$H_{\mathrm{BD}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J_R^*}{\partial p_C} F_{\mathrm{BD}}(p_C, \mu_R, t), \tag{S82}$$

其中，J_R^* 是针对如下最优奖励方案的 p_C 和 t 的最优成本函数

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R)^2}{2} dt. \tag{S83}$$

Sol求得$\frac{\partial H_{\mathrm{BD}}}{\partial \mu_R} = 0$ 后，我们知道最优奖励方案 μ_R^* 应满足

$$\mu_R^* = -\frac{\omega(k-2)(1-p_C)}{k(k-1)N^2 p_C} \frac{\partial J_R^*}{\partial p_C}. \tag{S84}$$

具有正向激励的动力系统对应的 HJB方程 [37–39] 可
写为

$$-\frac{\partial J_R^*}{\partial t} = H_{\text{BD}}(p_C, \mu_R^*, t). \tag{S85}$$

由于我们假设终止时间 t_f不固定，因此最优成本函数 $J_R^*(p_C, t)$与 t无关。因此，我们有

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S86}$$

于是得到

$$\frac{\partial J_R^*}{\partial p_C} = 0 \ \text{ or } \ \frac{\partial J_R^*}{\partial p_C} = -\frac{2N^2k(k-1)cp_C}{\omega(k-2)(1-p_C)}. \tag{S87}$$

由于 $\mu_R > 0$ 和 $p_C \in (0,1)$，我们有

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S88}$$

因此，只有 $\frac{\partial J_R^*}{\partial p_C} = -\frac{2N^2k}{(k-1)\omega(k-2)}cp_C(1-p_C)$ 成立。将该方程代入方程 (S84)，可得最优奖励方案：

$$\mu_R^* = 2c. \tag{S89}$$

采用最优奖励方案 μ_R^*后，动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)c}{k-1}p_C(1-p_C), \tag{S90}$$

其中，种群中的合作者初始比例记为 $p_C(0) = p_0$。求解
上述方程，得

$$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{BD}}t}}, \tag{S91}$$

其中 $\beta_{\text{BD}} = \frac{\omega k}{(k-2)}(k-1)c$。因此，最优奖励方案产生的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{BD}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S92}$$

如果我们求解惩罚的最优控制问题，就能得到负激励的最优协议

$$\mu_P^* = 2c, \tag{S93}$$

and

$$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{BD}}t}}. \tag{S94}$$

相应地，最优惩罚协议所需的累积成本为

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{BD}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S95}$$

因此，最优奖励与惩罚方案的累积成本差为

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\text{BD}}} \ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S96}$$

其中 $\mu_v^* = \mu_R^* = \mu_P^*$。与第 1.3 小节对方程 (S52) 的分析类似，我们还发现，$J_R^* > J_P^*$当 $\delta < p_0$ 时，以及 $J_R^* < J_P^*$当 $\delta > p_0$ 时。这意味着，在 BD 更新中，当 $\delta < p_0$ 时，执行最优惩罚方案相比执行最优奖励方案能够降低累积成本；而当 $\delta > p_0$时，结论则相反。数值计算和蒙特卡洛模拟也证实了这一结论，相关结果分别见图 4 和图 S5。

3. IM 更新

3.1. 正向激励

对于 IM 更新， [15, 16], 会随机选取一个焦点个体来更新其策略，该个体的 k 个邻居中有 k_C 个合作者和 k_D 个背叛者。如果焦点个体采用策略 D，那么一个 C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b - c + \mu_R)(k-1)q_{C|C} + (\mu_R - c)[(k-1)q_{D|C} + 1]\}, \tag{S97}$$

而 a D–邻居的适应度为

$$f_D = 1 - \omega + \omega\{b(k-1)q_{C|D} + 0 \cdot [(k-1)q_{D|D} + 1]\}. \tag{S98}$$

此外，焦点个体的适应度为

$$f_0 = 1 - \omega + \omega b k_C. \tag{S99}$$

由于焦点个体可以保留自己的策略，也可以按照与适应度成正比的概率模仿邻居的策略，因此焦点个体采用策略 C 的概率为

$$\Theta = \frac{k_C f_C}{k_C f_C + k_D f_D + f_0}. \tag{S100}$$

因此， p_C 以概率增加 $1/N$

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S101}$$

因此， CC–对的数量增加 k_C，进而 p_{CC} 以概率增加 $k_C/（kN/2）$

$$P(\Delta p_{CC}=\frac{2k_C}{kN})=p_D\binom{k}{k_C}(q_{C|D})^{k_C}(q_{D|D})^{k_D}\Theta. \tag{S102}$$

在另一种情况下，随机选中的焦点个体采用策略 C。此时，a C–邻居的适应度为

$$f_C=1-\omega+\omega\{(b-c+\mu_R)[(k-1)q_{C|C}+1]+(\mu_R-c)(k-1)q_{D|C}\}, \tag{S103}$$

而 a D–邻居的适应度为

$$f_D=1-\omega+\omega[(k-1)q_{C|D}+1]b. \tag{S104}$$

此外，焦点个体的适应度为

$$f_0=1-\omega+\omega[(b-c+\mu_R)k_C+(\mu_R-c)k_D]. \tag{S105}$$

焦点个体采用策略 D 的概率为

$$\Upsilon=\frac{k_Df_D}{k_Cf_C+k_Df_D+f_0}. \tag{S106}$$

因此， p_C 以 $1/N$ 的概率减少

$$P(\Delta p_C=-\frac{1}{N})=p_C\sum_{k_C=0}^{k}\binom{k}{k_C}(q_{C|C})^{k_C}(q_{D|C})^{k_D}\Upsilon. \tag{S107}$$

因此， CC–对的数量减少 k_C，从而 p_{CC} 以 $k_C/（kN/2）$ 的概率减少

$$P(\Delta p_{CC}=-\frac{2k_C}{kN})=p_C\binom{k}{k_C}(q_{C|C})^{k_C}(q_{D|C})^{k_D}\Upsilon. \tag{S108}$$

根据这些计算， p_C 的时间导数为

$$\begin{aligned}\frac{dp_C}{dt}&=\frac{E(\Delta p_C)}{\Delta t}=\frac{\frac{1}{N}P(\Delta p_C=\frac{1}{N})-\frac{1}{N}P(\Delta p_C=-\frac{1}{N})}{\frac{1}{N}}\\&=\frac{\omega kp_{CD}}{(k+1)^2}[2(\mu_R-c-b)+2(k-1)\xi_1+(k-1)(\mu_R-c)(q_{C|C}\\&+q_{D|D})+(k+1)^2q_{C|C}+q_{D|D}\xi_1]+o(\omega^2),\end{aligned} \tag{S109}$$

其中 $\xi_1=(b-c+\mu_R)q_{C|C}+(\mu_R-c)q_{D|C}-bq_{C|D}$。因此， p_{CC} 的时间导数为

$$\begin{aligned}\frac{dp_{CC}}{dt}&=\frac{E(\Delta p_{CC})}{\Delta t}=\frac{\sum_{k_C=0}^{k}\frac{2k_C}{kN}P(\Delta p_{CC}=\frac{2k_C}{kN})-\sum_{k_C=0}^{k}\frac{2k_C}{kN}P(\Delta p_{CC}=-\frac{2k_C}{kN})}{\frac{1}{N}}\\&=\frac{2p_{CD}}{k+1}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega).\end{aligned} \tag{S110}$$

此外，我们有

$$\frac{dq_{C|C}}{dt}=\frac{2p_{CD}}{(k+1)p_C}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega). \tag{S111}$$

因此，动力学方程可表示为

$$\begin{cases}\frac{dp_C}{dt}=\omega\Psi^R_{\mathrm{IM}}(p_C,q_{C|C})+o(\omega^2),\\ \frac{dq_{C|C}}{dt}=\Phi^R_{\mathrm{IM}}(p_C,q_{C|C})+o(\omega),\end{cases} \tag{S112}$$

其中

$$\begin{cases}\Psi^R_{\mathrm{IM}}(p_C,q_{C|C})=\frac{kp_{CD}}{(k+1)^2}[2(\mu_R-c-b)+2(k-1)\xi_1+(k-1)(\mu_R-c)(q_{C|C}\\ \qquad\qquad +q_{D|D})+(k+1)^2q_{C|C}+q_{D|D}\xi_1],\\ \Phi^R_{\mathrm{IM}}(p_C,q_{C|C})=\frac{2p_{CD}}{(k+1)p_C}[1+(k-1)(q_{C|D}-q_{C|C})].\end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很快，并可能迅速收敛到由 $\Phi^R_{\mathrm{IM}}(p_C,\ q_{C|C})=0$ 定义的根，此时 $t\to+\infty$。因此，我们得到

$$q_{C|C}=p_C+\frac{1}{k-1}(1-p_C). \tag{S113}$$

相应地，方程 (S112) 所描述的动力学方程变为

$$\frac{dp_C}{dt}=\frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)}p_C(1-p_C)+o(\omega^2), \tag{S114}$$

其有　作为两个不动点 $p_C=0$ 和 $p_C=1$。我们将函数 $F_{\mathrm{IM}}(p_C,\ \mu_R,\ t)$ 定义为 a　s

$$F_{\mathrm{IM}}(p_C,\mu_R,t)=\frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)}p_C(1-p_C)+o(\omega^2), \tag{S115}$$

以及 $F_{\mathrm{IM}}(p_C,\ \mu_R,\ t)$ 关于 p_C 的导数为

$$\frac{dF_{\mathrm{IM}}}{dp_C}=\frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)}(1-2p_C)+o(\omega^2). \tag{S116}$$

对于 $\mu_R>c-\frac{b}{k+2}$，我们有$\frac{dF_{\mathrm{IM}}}{dp_C}|_{p_C=1}=-\frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)}<0$以及$\frac{dF_{\mathrm{IM}}}{dp_C}|_{p_C=0}=\frac{\omega k^2(k-2)[b+(\mu_R-c)(k+2)]}{(k+1)^2(k-1)}>0$ 这意味着不动点 $p_C=1$ 是稳定的，而 $p_C=0$ 是不稳定的，即合作者胜过背叛者。特别地，当 $\mu_R=0$ 时，我们可以看到，对于$b/c>k+2$，不动点 $p_C=1$ 是稳定的，而 $p_C=0$ 是不稳定的。因此，我们得到了 IM 更新规则下合作演化的条件 $b/c>k+2$，这与文献 [15, 16] 中先前得到的结果一致。

3.2. 负激励

在本小节中，我们考虑将负激励引入采用 IM 更新的网络化囚徒困境博弈。根据这一规则，我们随机选择一个焦点个体来更新其策略，该个体的 k 个邻居中有 k_C 个合作者和 k_D 个背叛者。如果焦点个体采用策略 D，那么一个 C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b-c)(k-1)q_{C|C} - c[(k-1)q_{D|C} + 1]\}, \tag{S117}$$

且一个 D–邻居的适应度为

$$f_D = 1 - \omega + \omega\{(b-\mu_P)(k-1)q_{C|D} - \mu_P[(k-1)q_{D|D} + 1]\}. \tag{S118}$$

此外，焦点个体的适应度为

$$f_0 = 1 - \omega + \omega[(b-\mu_P)k_C - \mu_P k_D]. \tag{S119}$$

焦点个体采用策略 C 的概率由方程 (S100) 中的表达式 Θ 给出。因此，p_C 以概率增加 $1/N$

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S120}$$

此外，CC–对的数量增加 k_C，因此 p_{CC} 以概率 $(kN/2)$ 增加 $k_C/$

$$P(\Delta p_{CC} = \frac{2k_C}{kN}) = p_D \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \Theta. \tag{S121}$$

此外，我们考虑另一种情况，即随机选中的焦点个体采用策略 C。在这种情况下，一个 C–邻居的适应度为

$$f_C = 1 - \omega + \omega\{(b-c)[(k-1)q_{C|C} + 1] - c(k-1)q_{D|C}\}, \tag{S122}$$

而 a D–邻居的适应度为

$$f_D = 1 - \omega + \omega\{(b-\mu_P)[(k-1)q_{C|D} + 1] - \mu_P(k-1)q_{D|D}\}. \tag{S123}$$

此外，焦点个体的适应度为

$$f_0 = 1 - \omega + \omega[(b-c)k_C - ck_D]. \tag{S124}$$

焦点个体采用策略 D 的概率由方程（S106）中的表达式 Υ 给出。因此， p_C 以概率 $1/N$ 减少

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \Upsilon. \tag{S125}$$

并且， CC–对的数量减少 k_C ，因此 p_{CC} 以 $k_C/(kN/2)$的概率减少

$$P(\Delta p_{CC} = -\frac{2k_C}{kN}) = p_C \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \Upsilon. \tag{S126}$$

根据这些计算，我们得到 p_C 的时间导数为

$$\begin{aligned} \frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N} P(\Delta p_C = \frac{1}{N}) - \frac{1}{N} P(\Delta p_C = -\frac{1}{N})}{\frac{1}{N}} \\ &= \frac{\omega k p_{CD}}{(k+1)^2} \{2[(\mu_P - c - b) + (k-1)\xi_2] + (k-1)(\mu_P - c)(q_{C|C} + q_{D|D}) \\ &+ (k+1)^2 (q_{C|C} + q_{D|D})\xi_2\} + o(\omega^2), \end{aligned} \tag{S127}$$

其中， $\xi_2 = (b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}$。相应地， p_{CC} 的时间导数为

$$\begin{aligned} \frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = \frac{2k_C}{kN}) - \sum_{k_C=0}^{k} \frac{2k_C}{kN} P(\Delta p_{CC} = -\frac{2k_C}{kN})}{\frac{1}{N}} \\ &= \frac{2p_{CD}}{k+1}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \end{aligned} \tag{S128}$$

此外，我们有

$$\frac{dq_{C|C}}{dt} = \frac{2p_{CD}}{(k+1)p_C}[1 + (k-1)(q_{C|D} - q_{C|C})] + o(\omega). \tag{S129}$$

因此，动力学方程可表示为

$$\begin{cases} \frac{dp_C}{dt} = \omega \Psi^P_{\text{IM}}(p_C, q_{C|C}) + o(\omega^2), \\ \frac{dq_{C|C}}{dt} = \Phi^P_{\text{IM}}(p_C, q_{C|C}) + o(\omega), \end{cases} \tag{S130}$$

其中

$$\begin{cases} \Psi^P_{\text{IM}}(p_C, q_{C|C}) = \frac{k p_{CD}}{(k+1)^2} \{2[(\mu_P - c - b) + (k-1)\xi_2] + (k-1)(\mu_P - c)(q_{C|C} + q_{D|D}) \\ \qquad\qquad +(k+1)^2 (q_{C|C} + q_{D|D})\xi_2\}, \\ \Phi^P_{\text{IM}}(p_C, q_{C|C}) = \frac{2p_{CD}}{(k+1)p_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很快，并且可能迅速收敛到由 $\Phi^{P}_{\text{IM}}(p_C, q_{C|C}) = 0$ 定义的根，随着时间 $t \to +\infty$。因此，我们得到

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S131}$$

因此，方程 (S130) 所描述的动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega k^2(k-2)[b + (\mu_P - c)(k+2)]}{(k+1)^2(k-1)} p_C(1 - p_C) + o(\omega^2), \tag{S132}$$

其具有 两个不动点 $p_C = 0$和 $p_C = 1$。我们将函数 $F_{\text{IM}}(p_C,\ \mu_P,\ t)$ 定义为 s

$$F_{\text{IM}}(p_C, \mu_P, t) = \frac{\omega k^2(k-2)[b + (\mu_P - c)(k+2)]}{(k+1)^2(k-1)} p_C(1 - p_C) + o(\omega^2), \tag{S133}$$

$F_{\text{IM}}(p_C,\ \mu_P,\ t)$ 关于 p_C 的导数为

$$\frac{dF_{\text{IM}}}{dp_C} = \frac{\omega k^2(k-2)[b + (\mu_P - c)(k+2)]}{(k+1)^2(k-1)}(1 - 2p_C) + o(\omega^2). \tag{S134}$$

对于 $\mu_P > c - \frac{b}{k+2}$，我们有$\frac{dF_{\text{IM}}}{dp_C}|_{p_C=1} = \frac{\omega k^2(k-2)[b+(\mu_P-c)(k+2)]}{(k+1)^2(k-1)} < 0$，且 $\frac{dF_{\text{IM}}}{dp_C}|_{p_C=0} = \omega k^2(k-2)[b+(\mu_P-c)(k+2)]\ (k+1)^2(k-1) > 0$，这意味着不动点 $p_C = 1$ 稳定，而 $p_C = 0$不稳定，即合作者胜过背叛者。特别是，当 $\mu_P = 0$时，我们可以看到，对于 $b/c > k+2$，不动点 $p_C = 1$ 稳定，而 $p_C = 0$ 不稳定。因此，我们得到了 IM 更新规则下合作演化的条件$b/c > k+2$，与参考文献 [15, 16]中的结果一致。

3.3. 最优激励协议

在第 3.1 和 3.2 小节中，我们利用弱选择极限下的成对近似方法，从理论上得到了带有正激励或负激励的动力学方程，其形式为

$$\frac{dp_C}{dt} = F_{\text{IM}}(p_C, \mu_v, t) = \frac{\omega k^2(k-2)[b + (\mu_v - c)(k+2)]}{(k+1)^2(k-1)} p_C(1 - p_C) + o(\omega^2). \tag{S135}$$

该动力学方程有两个均衡点，分别为 $p_C = 0$ 和 $p_C = 1$。如果 $\mu_v > c - \frac{b}{k+2}$，前者不稳定，后者稳定，这意味着长期来看合作将得到促进。此外，为了确定最优奖励与惩罚协议，我们现在采用 HJB方程方法。

该控制问题的哈密顿函数为

$$H_{\text{IM}}(p_C, \mu_R, t) = \frac{(kNp_C\mu_R)^2}{2} + \frac{\partial J^*_R}{\partial p_C} F_{\text{IM}}(p_C, \mu_R, t), \tag{S136}$$

其中 J_R^* 是 t 给定 a 时， p_C 和 t 在最优奖励方案下的最优成本函数 s

$$J_R^* = \int_0^{t_f} \frac{(kNp_C\mu_R^*)^2}{2} dt. \tag{S137}$$

求解 $\frac{\partial H_{\text{IM}}}{\partial \mu_R} = 0$，可知最优奖励方案$\mu_R^*$ 应满足

$$\mu_R^* = -\frac{\omega(k-2)(k+2)(1-p_C)}{N^2(k-1)(k+1)^2 p_C} \frac{\partial J_R^*}{\partial p_C}. \tag{S138}$$

HJB方程可写为

$$-\frac{\partial J_R^*}{\partial t} = H_{\text{IM}}(p_C, \mu_R^*, t). \tag{S139}$$

由于终止时间 t_f不固定，最优成本函数 $J_R^*(p_C, t)$ 与 t 无关。因此，我们有

$$\frac{\partial J_R^*}{\partial t} = 0. \tag{S140}$$

于是得到

$$\frac{\partial J_R^*}{\partial p_C} = 0 \ \text{ or } \ \frac{\partial J_R^*}{\partial p_C} = \frac{2N^2 p_C[b - c(k+2)](k+1)^2(k-1)}{\omega(k-2)(k+2)^2(1-p_C)}. \tag{S141}$$

由于 $\mu_R > 0$ 和 $p_C \in (0,1)$，我们有

$$\frac{\partial J_R^*}{\partial p_C} < 0. \tag{S142}$$

由公式（S142）可知，对于 $b/c \geq k+2$，该不等式显然成立。相反，我们考虑 $b/c < k+2$ 这种情况，因此只有 $\frac{\partial J_R^*}{\partial p_C} = \frac{2N^2 p_C[b-c(k+2)](k+1)^2(k-1)}{\omega(k-2)(k+2)^2(1-p_C)}$ 成立。将该方程代入公式（S138），可得最优奖励方案为

$$\mu_R^* = \frac{2[c(k+2) - b]}{k+2}. \tag{S143}$$

采用最优奖励方案 μ_R^*后，动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega k^2(k-2)[c(k+2) - b]}{(k+1)^2(k-1)} p_C(1 - p_C), \tag{S144}$$

其中，种群中的合作者初始比例记为 $p_0 = p_C(0)$。为求解

上述方程，我们有

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0} e^{-\beta_{\text{IM}} t}}, \tag{S145}$$

其中 $\beta_{\text{IM}} = \frac{\omega k^2}{(k-2)[c(k+2)-b]}(k+1)^2(k-1)$。因此，最优奖励方案产生的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{IM}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S146}$$

然后，我们求解正文中方程 (17) 所描述的惩罚最优控制问题。经过计算，我们分别得到最优惩罚方案 μ_P^* 以及其对应的解 p_C，如下所示

$$\mu_P^* = \frac{2[c(k+2)-b]}{k+2}, \tag{S147}$$

and

$$p_C = \frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\text{IM}}t}}. \tag{S148}$$

因此，最优惩罚方案产生的累积成本为

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{IM}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S149}$$

因此，最优奖励与惩罚方案之间的累积成本差为

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\text{IM}}}\ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S150}$$

其中 $\mu_v^* = \mu_R^* = \mu_P^*$。与方程 (S52) 类似，我们还发现，当 $\delta < p_)$ 时，$J_R^* > J_P^*$；但当 $\delta > p_0$ 时，$J_R^* < J_P^*$。这表明，对于 IM 更新，当 $\delta < p_0$ 时，执行最优惩罚方案所需的累积成本低于执行最优奖励方案；而当 $\delta > p_0$ 时，结论则相反。这些理论结果可由数值计算和蒙特卡洛模拟加以验证，分别如图 4 和图 S5 所示。

4. PC 更新

4.1. 正向激励

对于 PC 更新 [16]，我们随机选择一个焦点个体来更新其策略，该个体的 k 个邻居中有 k_C 个合作者和 k_D 个背叛者。如果焦点个体采用策略 D，那么该焦点个体的适应度为

$$f_D = 1-\omega+\omega\pi_0^D = 1-\omega+\omega[bk_C+0\cdot k_D], \tag{S151}$$

且一个 C–邻居的适应度为

$$f_C = 1-\omega+\omega\pi_C^D = 1-\omega+\omega\{(b-c+\mu_R)(k-1)q_{C|C}+(\mu_R-c)[(k-1)q_{D|C}+1]\}. \tag{S152}$$

其中 π_0^D 表示 是焦点个体的收益，而 π_C^D 表示一个 C–邻居的收益 .

由于焦点个体要么保留当前策略，要么以取决于收益差（即 $\pi_C^D - \pi_0^D$）的概率采用邻居的策略，因此，对于 $\omega \to 0$，焦点个体采用一个 C–邻居策略的概率为

$$\Lambda = \frac{1}{1+e^{-\omega(\pi_C^D-\pi_0^D)}} = \frac{1}{2} + \omega\frac{\pi_C^D - \pi_0^D}{4}. \tag{S153}$$

由于弱选择下 $f_C - f_D = \omega(\pi_C^D - \pi_0^D)$，我们进一步得到

$$\Lambda = \frac{1}{1+e^{-\omega(\pi_C^D-\pi_0^D)}} = \frac{1}{1+e^{-(f_C-f_D)}} = \frac{1}{2} + \frac{f_C - f_D}{4}. \tag{S154}$$

因此， p_C 以 $1/N$ 的概率增加

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda. \tag{S155}$$

因此， CC–对的数量增加 $(k-1)q_{C|D}+1$，而 p_{CC} 以 $[(k-1)q_{C|D}+1]/(kN/2)$ 的概率增加

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda. \tag{S156}$$

此外，我们考虑另一种情况：随机选中的焦点个体采用策略 C。在这种情况下，焦点个体的适应度为

$$f_C = 1 - \omega + \omega\pi_0^C = 1 - \omega + \omega[(b - c + \mu_R)k_C + (\mu_R - c)k_D], \tag{S157}$$

而一个 D–邻居的适应度为

$$f_D = 1 - \omega + \omega\pi_D^C = 1 - \omega + \omega[(k-1)q_{C|D} + 1]b, \tag{S158}$$

其中， π_0^C 表示焦点个体的收益， π_D^C 表示采用策略 D 的邻居的收益。当$\omega \to 0$时，焦点个体采用一个 D–邻居策略的概率为

$$\Omega = \frac{1}{1+e^{-\omega(\pi_D^C-\pi_0^C)}} = \frac{1}{2} + \omega\frac{\pi_D^C - \pi_0^C}{4} = \frac{1}{2} + \frac{f_D - f_C}{4}. \tag{S159}$$

因此， p_C 以概率减少 $1/N$

$$P(\Delta p_C = -\frac{1}{N}) = p_C \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|C})^{k_C} (q_{D|C})^{k_D} \frac{k_D}{k} \Omega. \tag{S160}$$

因此， CC对的数量减少$(k-1)q_{C|C}$ ，从而 p_{CC} 以概率增加$(k-1)q_{C|C}/(kN/2)$

$$P(\Delta p_{CC}=-\frac{(k-1)q_{C|C}}{kN/2})=p_C\sum_{k_C=0}^{k}\binom{k}{k_C}(q_{C|C})^{k_C}(q_{D|C})^{k_D}\frac{k_D}{k}\Omega. \tag{S161}$$

基于这 些计算，我们分别得到 p_C 和 p_{CC} 的时间导数:

$$\begin{aligned}\frac{dp_C}{dt}&=\frac{E(\Delta p_C)}{\Delta t}=\frac{\frac{1}{N}P(\Delta p_C=\frac{1}{N})-\frac{1}{N}P(\Delta p_C=-\frac{1}{N})}{\frac{1}{N}}\\&=\frac{\omega p_{CD}}{2}\{(\mu_R-c-b)+(k-1)[(b-c+\mu_R)q_{C|C}+(\mu_R-c)q_{D|C}-bq_{C|D}]\}+o(\omega^2).\end{aligned} \tag{S162}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt}&=\frac{E(\Delta p_{CC})}{\Delta t}=\frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC}=\frac{(k-1)q_{C|D}+1}{kN/2})-\frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC}=-\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}}\\&=\frac{1}{k}p_{CD}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega).\end{aligned} \tag{S163}$$

此外，我们有

$$\frac{dq_{C|C}}{dt}=\frac{p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})]+o(\omega). \tag{S164}$$

因此，动力学方程可表示为

$$\begin{cases}\frac{dp_C}{dt}=\omega\Psi^R_{\text{PC}}(p_C,q_{C|C})+o(\omega^2),\\ \frac{dq_{C|C}}{dt}=\Phi^R_{\text{PC}}(p_C,q_{C|C})+o(\omega),\end{cases} \tag{S165}$$

其中

$$\begin{cases}\Psi^R_{\text{PC}}(p_C,q_{C|C})=\frac{p_{CD}}{2}\{(\mu_R-c-b)+(k-1)[(b-c+\mu_R)q_{C|C}+(\mu_R-c)q_{D|C}-bq_{C|D}]\}\\ \Phi^R_{\text{PC}}(p_C,q_{C|C})=\frac{p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})].\end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很大，并且可能迅速收敛到由 $\Phi^R_{\text{PC}}(p_C,\ q_{C|C})=0$ 定义的根，随着时间 $t\to+\infty$。因此，我们得到

$$q_{C|C}=p_C+\frac{1}{k-1}(1-p_C). \tag{S166}$$

因此，由公式（S165）描述的动力学方程变为

$$\frac{dp_C}{dt}=\frac{\omega k(k-2)(\mu_R-c)}{2(k-1)}p_C(1-p_C)+o(\omega^2), \tag{S167}$$

其具有两个不动点 $p_C = 0$ 和 $p_C = 1$。我们将函数 $F_{\text{PC}}(p_C,\ \mu_R,\ t)$ 定义为

$$F_{\text{PC}}(p_C, \mu_R, t) = \frac{\omega k(k-2)(\mu_R - c)}{2(k-1)} p_C(1-p_C) + o(\omega^2), \tag{S168}$$

以及 $F_{\text{PC}}(p_C,\ \mu_R,\ t)$ 关于 p_C 的导数为

$$\frac{dF_{\text{PC}}}{dp_C} = \frac{\omega k(k-2)(\mu_R - c)}{2(k-1)}(1-2p_C) + o(\omega^2). \tag{S169}$$

对于 $\mu_R > c$，我们有 $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_R-c)}{2(k-1)} < 0$ 以及 $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=0} = \frac{\omega k(k-2)(\mu_R-c)}{2(k-1)} > 0$，这意味着不动点 $p_C = 1$ 是稳定的，而 $p_C = 0$ 是不稳定的，即合作者胜过背叛者。特别地，当 $\mu_R = 0$ 时，我们可以看到不动点 $p_C = 0$ 始终稳定，而 $p_C = 1$ 不稳定，这意味着正如先前的研究[16]所观察到的，合作永远不会出现。

4.2. 负激励

在本小节中，我们考察惩罚在 PC 更新中如何发挥作用。根据这一规则，我们随机选择一个焦点个体来更新其策略；该个体的 k 个邻居中有 k_C 个合作者和 k_D 个背叛者。如果焦点个体采用策略 D，那么该焦点个体的适应度为

$$f_D = 1 - \omega + \omega\pi_0^D = 1 - \omega + \omega[(b-\mu_P)k_C - \mu_P k_D], \tag{S170}$$

而 a C–邻居的适应度为

$$f_C = 1 - \omega + \omega\pi_C^D = 1 - \omega + \omega\{(b-c)(k-1)q_{C|C} - c[(k-1)q_{D|C} + 1]\}, \tag{S171}$$

其中，π_0^D表示焦点个体的收益，而 π_C^D表示 C–邻居的收益。

焦点个体采用某个 C-邻居策略的概率由方程 (S154) 中的表达式 Λ 给出。因此，p_C 以 $1/N$ 的概率增加

$$P(\Delta p_C = \frac{1}{N}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda, \tag{S172}$$

CC–配对的数量增加了 $(k-1)q_{C|D} + 1$，因此 p_{CC} 以概率 $[(k-1)q_{C|D} + 1]/(kN/2)$ 增加

$$P(\Delta p_{CC} = \frac{(k-1)q_{C|D}+1}{kN/2}) = p_D \sum_{k_C=0}^{k} \binom{k}{k_C} (q_{C|D})^{k_C} (q_{D|D})^{k_D} \frac{k_C}{k} \Lambda. \tag{S173}$$

在另一种情况下，随机选中的焦点个体采用策略 C。此时，焦点个体的适应度为

$$f_C = 1-\omega+\omega\pi_0^C = 1-\omega+\omega[(b-c)k_C - ck_D], \tag{S174}$$

而 a D–邻居的适应度为

$$f_D = 1-\omega+\omega\pi_D^C = 1-\omega+\omega\{(b-\mu_P)[(k-1)q_{C|D}+1]-\mu_P(k-1)q_{D|D}\}, \tag{S175}$$

其中， π_0^C表示焦点个体的收益， π_D^C 表示策略为 D的邻居的收益。焦点个体采用 a D–邻居策略的概率由方程 (S159) 中的表达式 Ω 给出。因此， p_C 以概率 $1/N$ 减少

$$P(\Delta p_C = -\frac{1}{N}) = p_C\sum_{k_C=0}^{k}\binom{k}{k_C}(q_{C|C})^{k_C}(q_{D|C})^{k_D}\frac{k_D}{k}\Omega. \tag{S176}$$

而且， CC–对的数量以概率减少 $(k-1)q_{C|C}$ ，因此 p_{CC} 以概率增加 $(k-1)q_{C|C}/(kN/2)$

$$P(\Delta p_{CC} = -\frac{(k-1)q_{C|C}}{kN/2}) = p_C\sum_{k_C=0}^{k}\binom{k}{k_C}(q_{C|C})^{k_C}(q_{D|C})^{k_D}\frac{k_D}{k}\Omega. \tag{S177}$$

基于这 些计算，我们分别得到 p_C 和 p_{CC} 的时间导数:

$$\begin{aligned}\frac{dp_C}{dt} &= \frac{E(\Delta p_C)}{\Delta t} = \frac{\frac{1}{N}P(\Delta p_C=\frac{1}{N})-\frac{1}{N}P(\Delta p_C=-\frac{1}{N})}{\frac{1}{N}}\\ &= \frac{\omega p_{CD}}{2}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\} + o(\omega^2).\end{aligned} \tag{S178}$$

and

$$\begin{aligned}\frac{dp_{CC}}{dt} &= \frac{E(\Delta p_{CC})}{\Delta t} = \frac{\frac{(k-1)q_{C|D}+1}{kN/2}P(\Delta p_{CC}=\frac{(k-1)q_{C|D}+1}{kN/2})-\frac{(k-1)q_{C|C}}{kN/2}P(\Delta p_{CC}=-\frac{(k-1)q_{C|C}}{kN/2})}{\frac{1}{N}}\\ &= \frac{p_{CD}}{k}[1+(k-1)(q_{C|D}-q_{C|C})] + o(\omega).\end{aligned} \tag{S179}$$

此外，我们有

$$\frac{dq_{C|C}}{dt} = \frac{p_{CD}}{kp_C}[1+(k-1)(q_{C|D}-q_{C|C})] + o(\omega). \tag{S180}$$

因此，动力学方程可表示为

$$\begin{cases}\frac{dp_C}{dt} = \omega\Psi_{\mathrm{PC}}^P(p_C, q_{C|C}) + o(\omega^2),\\ \frac{dq_{C|C}}{dt} = \Phi_{\mathrm{PC}}^P(p_C, q_{C|C}) + o(\omega),\end{cases} \tag{S181}$$

其中

$$\begin{cases} \Psi^P_{\text{PC}}(p_C, q_{C|C}) = \frac{p_{CD}}{2}\{\mu_P - c - b + (k-1)[(b-c)q_{C|C} - cq_{D|C} + (\mu_P - b)q_{C|D} + \mu_P q_{D|D}]\}, \\ \Phi^P_{\text{PC}}(p_C, q_{C|C}) = \frac{p_{CD}}{kp_C}[1 + (k-1)(q_{C|D} - q_{C|C})]. \end{cases}$$

在弱选择下，$q_{C|C}$ 的速度可能很大，并可能迅速收敛到由 $\Phi^P_{\text{PC}}(p_C,\ q_{C|C}) = 0$ 定义的根，随着时间 $t \to +\infty$。因此，我们得到

$$q_{C|C} = p_C + \frac{1}{k-1}(1 - p_C). \tag{S182}$$

相应地，方程 (S181) 所描述的动力学方程变为

$$\frac{dp_C}{dt} = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} p_C(1 - p_C) + o(\omega^2), \tag{S183}$$

它有两个不动点 $p_C = 0$和 $p_C = 1$。我们将函数 $F_{\text{PC}}(p_C,\ \mu_P,\ t)$ 定义为

$$F_{\text{PC}}(p_C, \mu_P, t) = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} p_C(1 - p_C) + o(\omega^2), \tag{S184}$$

$F_{\text{PC}}(p_C,\ \mu_P,\ t)$ 关于 p_C的导数为

$$\frac{dF_{\text{PC}}}{dp_C} = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)}(1 - 2p_C) + o(\omega^2). \tag{S185}$$

对于 $\mu_P > c$，我们有 $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=1} = -\frac{\omega k(k-2)(\mu_P - c)}{2(k-1)}() < 0$ 和 $\frac{dF_{\text{PC}}}{dp_C}|_{p_C=0} = \frac{\omega k(k-2)(\mu_P - c)}{2(k-1)} > 0$，这意味着不动点 $p_C = 1$ 稳定，而 $p_C = 0$ 不稳定，即合作者胜过背叛者。特别是，当 $\mu_P = 0$时，我们可以看出，不动点 $p_C = 0$ 始终稳定，而$p_C = 1$不稳定，这意味着合作永远无法涌现，正如参考文献 [16]所得出的结论。

4.3. 最优激励协议

通过成对近似方法，在弱选择极限下，采用PC更新规则的动力学方程为

$$\frac{dp_C}{dt} = F_{\text{PC}}(p_C, \mu_v, t) = \frac{\omega k(k-2)(\mu_v - c)}{2(k-1)} p_C(1 - p_C) + o(\omega^2). \tag{S186}$$

该动力学方程有两个均衡点，分别为 $p_C = 0$ 和 $p_C = 1$。如果 $\mu_v > c$，前者不稳定，后者稳定，因此从长远来看，合作将得到促进。此外，为了探究最优奖励与惩罚协议，我们现在采用 HJB方程的方法。

在奖励的情况下，我们将哈密顿函数 $H_{\mathrm{PC}}(p_C,\ \mu_R,\ t)$ 定义为

$$H_{\mathrm{PC}}(p_C,\mu_R,t)=\frac{(kNp_C\mu_R)^2}{2}+\frac{\partial J_R^*}{\partial p_C}F_{\mathrm{PC}}(p_C,\mu_R,t), \tag{S187}$$

其中 J_R^* 是 t 给定最优奖励方案时，p_C 和 t 的最优成本函数 s

$$J_R^*=\int_0^{t_f}\frac{(kNp_C\mu_R^*)^2}{2}dt. \tag{S188}$$

求解 $\frac{\partial H_{\mathrm{PC}}}{\partial \mu_R}=0$，可知最优奖励方案 μ_R^* 应满足

$$\mu_R^*=-\frac{\omega(k-2)(1-p_C)}{2N^2k(k-1)p_C}\frac{\partial J_R^*}{\partial p_C}. \tag{S189}$$

HJB方程可写为

$$-\frac{\partial J_R^*}{\partial t}=H_{\mathrm{PC}}(p_C,\mu_R^*,t). \tag{S190}$$

由于终止时间 t_f不固定，最优成本函数 $J_R^*(p_C,\ t)$ 与 t无关。因此，我们有

$$\frac{\partial J_R^*}{\partial t}=0. \tag{S191}$$

由此可得

$$\frac{\partial J_R^*}{\partial p_C}=0\ \text{ or }\ \frac{\partial J_R^*}{\partial p_C}=-\frac{4N^2k(k-1)cp_C}{\omega(k-2)(1-p_C)}. \tag{S192}$$

由于 $\mu_R>0$ 和 $p_C\in(0,\ 1)$，我们有

$$\frac{\partial J_R^*}{\partial p_C}<0. \tag{S193}$$

因此，只有 $\frac{\partial J_R^*}{\partial p_C}=-\frac{4N^2k}{(k-1)\omega\ k-2}$ ^{cp_C} $_{()(1-p_C)}$ 成立。将此方程代入方程（S189），我们得到最优奖励方案 μ_R^* 如下

$$\mu_R^*=2c. \tag{S194}$$

采用最优奖励方案μ_R^*后，动力学方程因此变为

$$\frac{dp_C}{dt}=\frac{\omega k(k-2)c}{2(k-1)}p_C(1-p_C), \tag{S195}$$

其中，种群中合作者的初始比例记为 $p_0=p_C(0)$。通过求解该方程，我们得到

$$p_C=\frac{1}{1+\frac{1-p_0}{p_0}e^{-\beta_{\mathrm{PC}}t}}, \tag{S196}$$

其中 $\beta_{\text{PC}} = \frac{\omega k}{(k-2)2^{k-1}}^{c}()$。因此，最优奖励方案产生的累积成本为

$$J_R^* = \frac{(kN\mu_R^*)^2}{2\beta_{\text{PC}}}[p_0 - 1 + \delta + \ln(\frac{1-p_0}{\delta})]. \tag{S197}$$

对于惩罚，我们分别得到最优协议 μ_P^* 及其对应的解 p_C，如下：

$$\mu_P^* = 2c \tag{S198}$$

and

$$p_C = \frac{1}{1 + \frac{1-p_0}{p_0}e^{-\beta_{\text{PC}}t}}. \tag{S199}$$

因此，最优惩罚方案产生的累积成本为

$$J_P^* = \frac{(kN\mu_P^*)^2}{2\beta_{\text{PC}}}[p_0 - 1 + \delta + \ln(\frac{1-\delta}{p_0})]. \tag{S200}$$

因此，累积成本值之差为

$$J_R^* - J_P^* = \frac{(kN\mu_v^*)^2}{2\beta_{\text{PC}}}\ln[\frac{p_0(1-p_0)}{\delta(1-\delta)}], \tag{S201}$$

其中 $\mu_v^* = \mu_R^* = \mu_P^*$。与方程 (S52) 类似，我们还发现，当 $\delta < p_0$ 时，$J_R^* > J_P^*$；但当 $\delta > p_0$ 时，我们有 $J_R^* < J_P^*$。这表明，对于 PC 更新，当 $\delta < p_0$ 时，采用最优惩罚比采用最优奖励方案成本更低；而当 $\delta > p_0$ 时，结论则相反。这些理论结果可由数值计算和蒙特卡洛模拟加以验证，分别如图 4 和图 S5 所示。

补充图

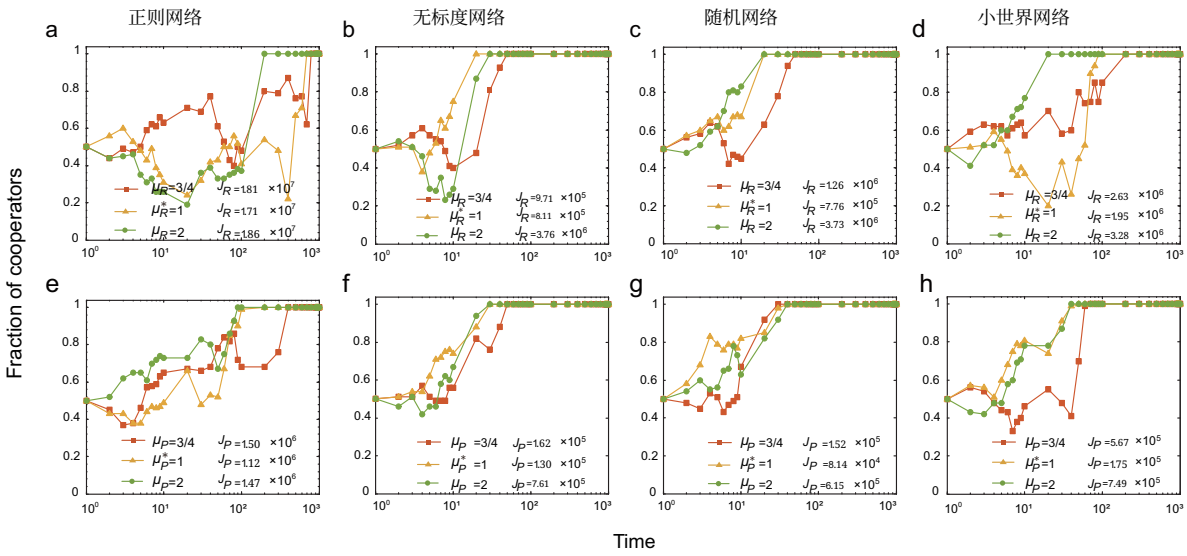

图 1：在 DB 更新下，四种不同网络中三种不同激励方案下合作者比例的时间演化。 所采用的方案标注在图例中，其中最优方案以 $*$表示。我们还绘制了各激励方案的累积成本值。奖励（惩罚）的蒙特卡洛模拟结果分别显示在上排（下排）。参数：$N = 100$、 $L = 10$、 $b = 2$、 $c = 1$、 $\delta = 0.01$、 $\omega = 0.01$和 $p_0 = 0.5$。为便于比较，所有图的平均度均设为 4。

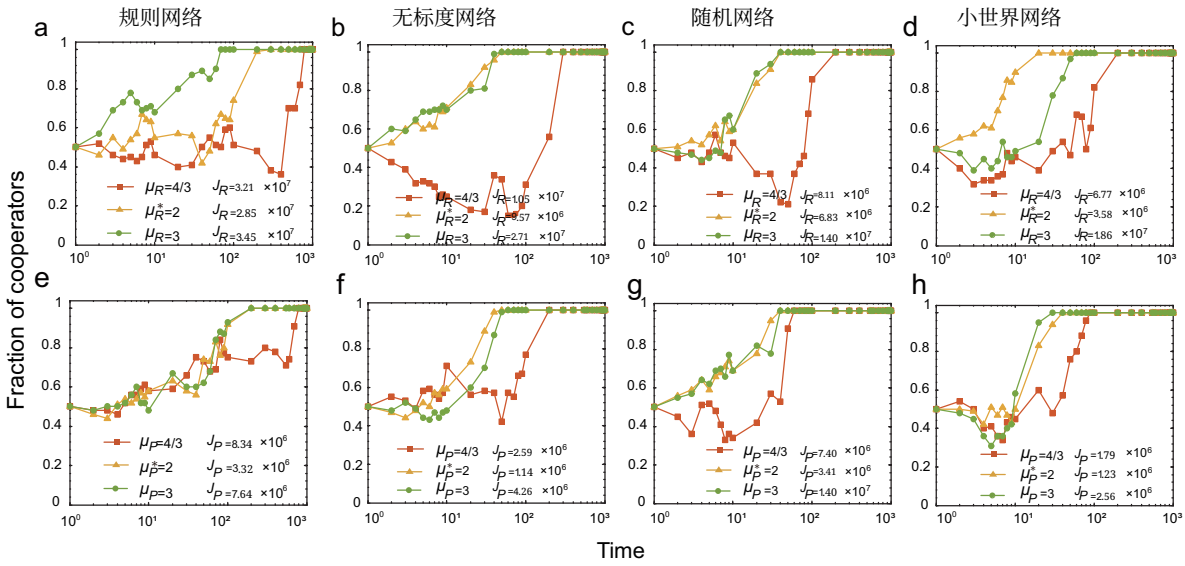

图 2：BD 更新下，四种不同网络中三种不同激励方案的合作者比例随时间的演化。应用的方案标记在图例中，其中最优方案由 *表示。我们还绘制了各激励方案的累积成本值。奖励（惩罚）的蒙特卡洛模拟结果分别显示在上排（下排）。其他参数值与图 S1 中相同。

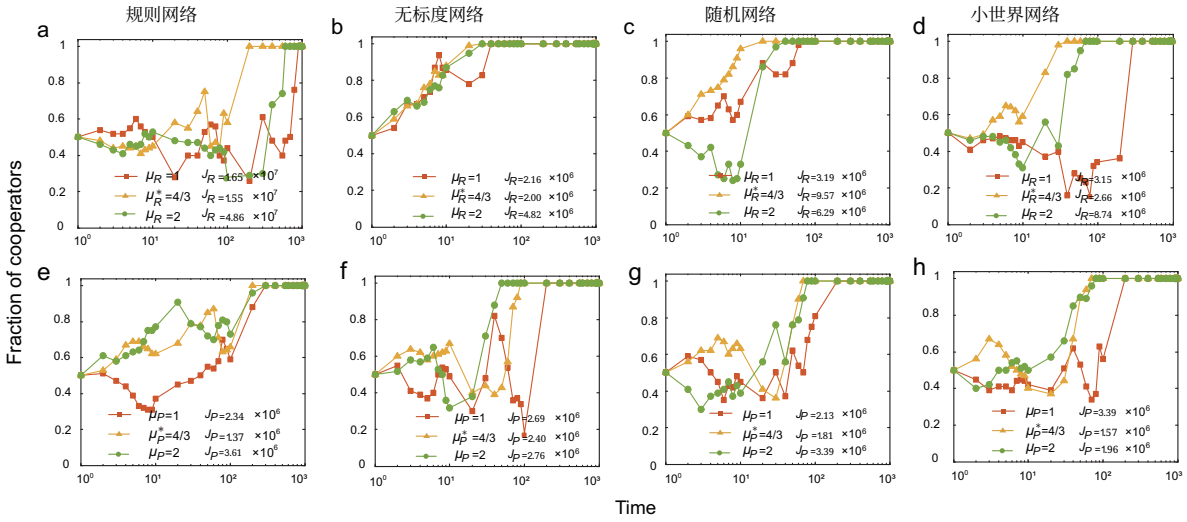

图 3：在 IM 更新下，四种不同网络中三种不同激励方案对应的合作者比例随时间的演化。所采用的方案由图例标出，其中最优方案以 *表示。我们还绘制了各激励方案的累积成本值。奖励（惩罚）的蒙特卡洛模拟结果分别显示在上（下）排。其他参数值与图 S1 中相同。

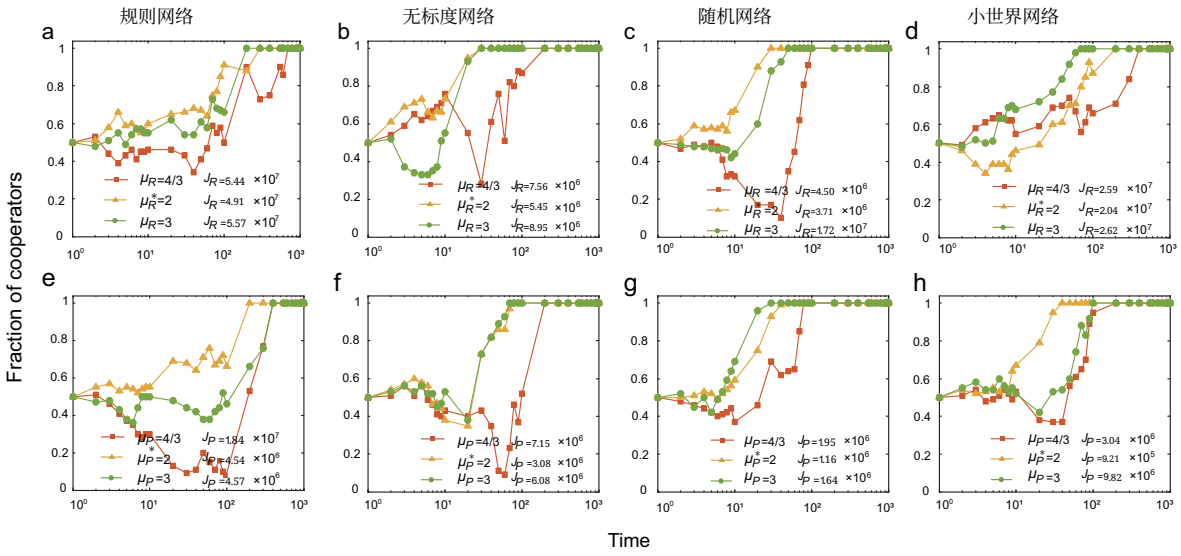

图 4：PC 更新下，四种不同网络中三种不同激励方案对应的合作者比例随时间的演化。所采用的方案由图例标示，其中最优方案以 *标出。我们还绘制了每种激励方案的累积成本值。奖励（惩罚）的蒙特卡洛模拟结果分别显示在上（下）排。其他参数值与图 S1 中相同。

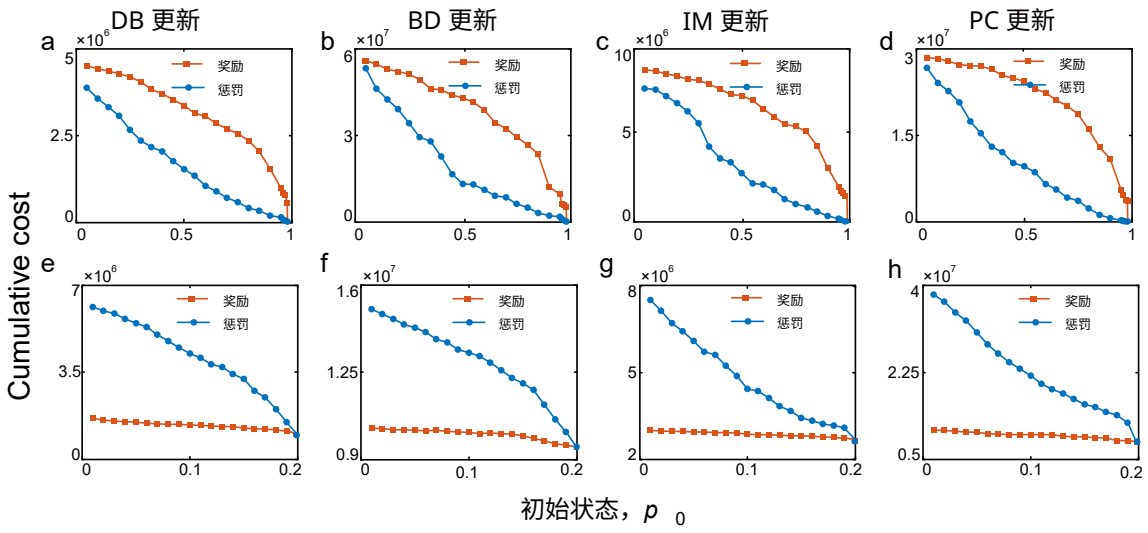

图5：在最优奖励与惩罚方案下，达到预期终止状态所需的累积成本与 p_0 合作者初始比例之间的关系。每列面板表示一种策略更新规则，如图所示。顶行表示蒙特卡洛模拟的结果，即在度为 k 的正则网络上，在$p_0 > \delta = 0.01$条件下进行 200 次独立模拟并取平均值；底行表示在度为 k 的正则网络上，在 $p_0 < \delta = 0.2$条件下进行 200 次独立蒙特卡洛模拟并取平均值所得的结果。其他参数： $N = 100$、 $L = 10$、 $b = 2$、 $c = 1$、 $\omega = 0.01$ 和 $k = 4$。