SCIENCE CHINA
Information Sciences

• Supplementary File •

Optimal tax-subsidy incentive for population games based on mean field approximation

Yingying Chai[1], Zhenyu Wu[2], Yuhu Wu[1*] & Shuting Le[3]

[1]*School of Control Science and Engineering, Dalian University of Technology, Dalian 116024, China*
[2]*School of Innovation and Entrepreneurship, Dalian University of Technology, Dalian 116024, China*
[3]*School of Electronic Information and Engineering, Suzhou University of Science and Technology, Suzhou 215000, China*

Appendix A Derivations of Equation (1)

The revised payoff of the prisoner's dilemma with the incentive parameter u^n is

$$\tilde{\pi}_s(x^n,u^n)=\begin{cases}\pi_s(x^n)-x^nu^n+(1-\alpha)u^n, & s=C,\\ \pi_s(x^n)-x^nu^n, & s=D,\end{cases} \tag{A1}$$

where $\pi_s(x^n)$ is the payoff for s-strategists without incentive, and x^n is the fraction of cooperators in the n-player population. In (A1), the second term denotes the capitation tax, and the third term is the subsidy treated as 0 when this term is non-existent.

The derivations, inspired by [1], are as follows. For the pairwise comparison update process, no more than one player modifies their strategy in each step, which can be expressed as

$$\mathbb{P}\mathrm{rob}\big(|x^n(\tau+1)-x^n(\tau)|>1/n\big|x^n(\tau)\big)\equiv 0,\quad \tau\in\mathbb{N}.$$

Then the analysis reduces to the case where $|x^n(\tau+1)-x^n(\tau)|\leqslant 1/n$, which encompasses the following three scenarios.

The first scenario is that a defector is chosen to mimic one of the $x^n\cdot n$ cooperators, resulting in x^n increased by $1/n$. The probability for this situation with the incentive parameter u^n is denoted as

$$P^+(x^n,u^n)=(1-x^n)\cdot\frac{nx^n}{n-1}\cdot\frac{1}{1+\exp\big(-\omega(\tilde{\pi}_C(x^n,u^n)-\tilde{\pi}_D(x^n,u^n))\big)}, \tag{A2}$$

where the first term on the right-hand side measures the probability of choosing defectors, the subsequent term represents the probability of choosing the cooperator as exemplar, and the terminal term denotes the probability of successful imitation.

The second scenario is that a cooperator is chosen to copy defectors when the fraction of cooperators is x^n at step τ, which leads to the population state decreasing from x^n to x^n-1/n. The probability relevant to this scenario with u^n is

$$P^-(x^n,u^n)=x^n\cdot\frac{n(1-x^n)}{n-1}\cdot\frac{1}{1+\exp\big(-\omega(\tilde{\pi}_D(x^n,u^n)-\tilde{\pi}_C(x^n,u^n))\big)}. \tag{A3}$$

The last scenario, in which no player changes strategies, has the probability expressed as

$$P^0(x^n,u^n)=1-P^+(x^n,u^n)-P^-(x^n,u^n). \tag{A4}$$

Let $F(x^n,\tau)$ denote the probability for the population state x^n at step τ. Based on Chapman-Kolmogorov equation in [2], $F(x^n,\tau+1)$ satisfies:

$$\begin{aligned}F(x^n,\tau+1)=&P^+(x^n-1/n,u^n)F(x^n-1/n,\tau)+P^-(x^n+1/n,u^n)F(x^n+1/n,\tau)+P^0(x^n,u^n)\big)F(x,\tau)\\ =&P^+(x^n-1/n,u^n)F(x^n-1/n,\tau)+P^-(x^n+1/n,u^n)F(x^n+1/n,\tau)+F(x^n,\tau)\\ &-P^+(x^n,u^n)F(x^n,\tau)-P^+(x^n,u^n)F(x^n,\tau),\end{aligned} \tag{A5}$$

where the second equality follows from (A4). Introducing the re-scaled time $t=\tau\omega/n$, the Taylor expansions of the probability $\rho(x,t)=F(x^n,\tau)$ for the re-scaled process $\{x(t)\}_{t\in\mathbb{R}_{\geqslant 0}}$ with the incentive policy $\{u(t):u(t)=u^n(\lfloor tn/\omega\rfloor)\}_{t\in\mathbb{R}_{\geqslant 0}}$ is given by

$$\rho(x,t+\Delta t)=\rho(x,t)+\partial_t\rho(x,t)\Delta t+o(\Delta t), \tag{A6}$$

* Corresponding author (email: wuyuhu@dlut.edu.cn)

where $\partial_t = \partial/\partial t$, and $\Delta t = \omega/n$.

Analogously, the Taylor expansions of $P^+\big(x-1/n,u\big)\rho(x-1/n,t), P^-\big(x+1/n,u\big)\rho(x+1/n,t)$ in terms of x can be, respectively, calculated as:

$$\begin{aligned}&P^+(x-1/n,u)\rho(x-1/n,t)\\=&P^+(x,u)\rho(x,t)-\frac{1}{n}\partial_x\big(P^+(x,u)\rho(x,t)\big)+\frac{1}{2n^2}\partial_{xx}\big(P^+(x,u)\rho(x,t)\big)+o\big(n^{-2}\big),\end{aligned}\tag{A7}$$

$$\begin{aligned}&P^-(x+1/n,u)\rho(x+1/nx,t)\\=&P^-(x,u)\rho(x,t)+\frac{1}{n}\partial_x\big(P^-(x,u)\rho(x,t)\big)+\frac{1}{2n^2}\partial_{xx}\big(P^-(x,u)\rho(x,t)\big)+o\big(n^{-2}\big),\end{aligned}\tag{A8}$$

where $\partial_x = \partial/\partial x$, and $\partial_{xx} = \partial^2/\partial x^2$. Eliminating $\rho(x,t)$ in (A5) and substituting the expressions given by (A6), (A7), and (A8), we get

$$\begin{aligned}\partial_t\rho(x,t)\Delta t+o(\Delta t)=&-\frac{1}{n}\partial_x P^+(x,u)\,\rho(x,t)+\frac{1}{n}\partial_x P^-(x,u)\,\rho(x,t)\\&+\frac{1}{2n^2}\partial_{xx}P^+(x,u)\,\rho(x,t)+\frac{1}{2n^2}\partial_{xx}P^-(x,u)\,\rho(x,t)+o\big(n^{-2}\big).\end{aligned}\tag{A9}$$

Let us now divide both sides of (A9) by Δt. For $n \gg 1$, we first neglect higher-order terms in n^{-1} $(= \Delta t)$ and then substitute (A2) and (A3) into (A9). On this basis, the Fokker-Planck equation for $\rho(x,t)$ is derived as

$$\partial_t\rho(x,t) = -\partial_x\mu(x,u)\rho(x,t)+\frac{1}{2}\partial_{xx}\sigma^2(x,u)\rho(x,t),$$

where

$$\mu(x,u)=\frac{1}{\omega}\frac{n}{n-1}x(1-x)\tanh\big(\frac{\omega}{2}\big(\tilde{\pi}_C(x,u)-\tilde{\pi}_D(x,u)\big)\big)\tag{A10}$$

is the drift coefficient and

$$\sigma^2(x,u)=\frac{1}{\omega n}x(1-x)$$

is the diffusion coefficient. Using Itô calculus, the stochastic differential equation is given by

$$dx=\mu(x,u)dt+\sigma(x,u)dB_t.\tag{A11}$$

where B_t is the Wiener process. For the weak selection $\omega \ll 1$, we get

$$\tanh\big(\frac{\omega}{2}\big(\tilde{\pi}_C(x,u)-\tilde{\pi}_D(x,u)\big)\big)=\frac{\omega}{2}\big(\tilde{\pi}_C(x,u)-\tilde{\pi}_D(x,u)\big)+o(\omega^2).\tag{A12}$$

By substituting (A12) into (A10), and subsequently neglecting higher-order terms in ω, the stochastic differential equation (A11) reduces to

$$dx=\frac{1}{2}\frac{n}{n-1}x(1-x)\big(\tilde{\pi}_C(x,u)-\tilde{\pi}_D(x,u)\big)dt+\sqrt{\frac{x(1-x)}{\omega n}}dB_t.$$

As $n\to\infty$, $\frac{n}{n-1}$ and $\sqrt{\frac{x(1-x)}{n}}$ converge to 1 and 0, respectively, leading to the deterministic equation

$$\dot{x}=\frac{1}{2}x(1-x)\big(\tilde{\pi}_C(x,u)-\tilde{\pi}_D(x,u)\big).\tag{A13}$$

Substituting (A1) into (A13) completes the derivation.

Appendix B Proof of Theorem 1

Proof. For the case where $x_1 = x_0$, leading the terminal time $T = t_0$, which leads

$$J(u)=0,\quad u\in U,$$

where $U = [0, b]$ is the admissible incentive parameter set. Obviously, $u^* = 0$ is the optimal policy.

For the other cases, introduce the value function as

$$V(x,t)=\inf_{u_{[t,T]}\in U_{[t,T]}}\int_t^T\frac{1}{2}x^2u^2ds,\quad t\in[t_0,T],$$

where $u_{[t,T]}$ is the incentive policy restricted to the interval $[t,T]$. Based on the Hamilton-Jacobi-Bellman equation in [3], the value function $V(x,t)$ follows:

$$-\partial_t V(x,t) = \inf_{u\in U}\Big\{\frac{1}{2}x^2u^2 + \partial_x V(x,t)\cdot\frac{1}{2}x(1-x)\big[(1-\alpha)u-c\big]\Big\}. \tag{B1}$$

Expressing the above equation in terms of the Hamiltonian function, it is straightforward to verify that (B1) is equivalent to

$$-\partial_t V(x,t) = \inf_{u(t)\in U} H(x,u,\partial_x V(t,x)),$$

where the Hamiltonian function is defined as

$$\begin{aligned} H(x,u,p) &= L(x,u) + p\cdot f(x,u)\\ &= \frac{1}{2}x^2u^2 + p\cdot\frac{1}{2}x\,(1-x)\,\big[\,(1-\alpha)\,u-c\big], \end{aligned}$$

with canonical equations

$$\begin{aligned} \partial_p H &= \dot{x},\\ \partial_x H &= -\dot{p}, \end{aligned} \tag{B2}$$

and a stationary condition

$$H\big(x(T),u(T),p(T)\big) = 0.$$

For the optimal state $x^*(t)$ driven by optimal control $u^*_{[t_0,t]}$, we obtain the following inequality:

$$H\big(x^*(t),u^*(t),\partial_x V(t,x^*(t))\big) \leqslant H\big(x^*(t),u,\partial_x V(t,x^*(t))\big),\ u\in U,$$

which yields that the optimal policy $u^*(t)$ is given by

$$u^*(t) = \begin{cases} \widetilde{u}(t), & \widetilde{u}(t)\in[0,b],\\ 0, & \widetilde{u}(t)<0,\\ b, & \widetilde{u}(t)>b, \end{cases}$$

where

$$\begin{aligned} \widetilde{u} &= \arg\min_u H(x,u,p)\\ &= \Big\{u : \frac{\partial H(x,u,p)}{\partial u} = 0\Big\}, \end{aligned} \tag{B3}$$

enables the convex function $H\big(x(t),u,\partial_x V(t,x(t))\big)$ to attain the minimum value.

By introducing (B2) and (B3), we obtain that for each $t\in[t_0,T]$

$$\begin{aligned} H\big(x(t),\widetilde{u},\partial_x V(t,x(t))\big) &= -\int_t^T \big(\partial_x H\cdot\dot{x} + \partial_u H\big|_{u=\widetilde{u}}\cdot\dot{u} + \partial_p H\cdot\dot{p}\big)dt + H\big(x(T),\widetilde{u},\partial_x V(T,x(T))\big)\\ &= H\big(x(T),\widetilde{u},\partial_x V(T,x(T))\big)\\ &= 0. \end{aligned} \tag{B4}$$

By combining (B3) and (B4), we derive

$$\widetilde{u}(t)\in\big\{0,\frac{2c}{1-\alpha}\big\}.$$

As a result, the optimal tax-subsidy incentive is characterized by $u^*\in\big\{0,\min\{\frac{2c}{1-\alpha},b\}\big\}$. The optimal population state trajectory x^*, governed by the dynamics $f(x^*,u^*)$, can be formulated as:

$$x^*(t) = x_0 + \frac{1}{2}\int_{t_0}^t x^*(1-x^*)\big[(1-\alpha)u^*-c\big]ds \implies x^*(t) = \big(1+\zeta e^{-\kappa(t-t_0)}\big)^{-1}, \tag{B5}$$

where $\kappa = \frac{1}{2}\left[(1-\alpha)\,u^*-c\right]$, and $\zeta = \frac{1-x_0}{x_0}$ are auxiliary parameters.

For the scenario where $x_1 < x_0$, the optimal state trajectory satisfies the following inequality:

$$\big(1+\zeta e^{-\kappa(T-t_0)}\big)^{-1} < x_0,\ T>t_0 \implies e^{-\kappa(T-t_0)} > 1,\ T>t_0,$$

which results in the condition $\kappa<0$. Consequently, the corresponding optimal policy for this scenario is $u^*=0$.

Conversely, for the case $x_1 > x_0$, the optimal policy satisfies $(1-\alpha)u^* - c > 0$, which implies

$$\begin{aligned} u^* &= \min\{\frac{2c}{1-\alpha}, b\} \\ &> \frac{c}{1-\alpha}. \end{aligned} \tag{B6}$$

Notably, a feasible u^* exists for $x_1 > x_0$ if and only if $\alpha \in [0, 1-\frac{c}{b})$; otherwise, no feasible policy exists to drive the population state toward the target state x_1.

The terminal condition $x^*(T^*) = x_1$ enables the optimal terminal time T^* to be represented by

$$T^* = \kappa^{-1} \ln\left(\zeta \cdot \eta^{-1}\right) + t_0. \tag{B7}$$

where $\eta = \frac{1-x_1}{x_1}$.

By substituting (B5) and (B7), the minimum value of the cost function $J(u^*)$ is calculated as

$$\begin{aligned} J(u^*) &= \int_{t_0}^{T^*} \frac{1}{2}\big(x^*(t)u^*(t)\big)^2 dt \\ &= \frac{1}{2}u^{*2}\kappa^{-1}\left(\ln(1+\eta^{-1}) + \frac{1}{1+\eta^{-1}} - \ln(1+\zeta^{-1}) - \frac{1}{1+\zeta^{-1}}\right). \end{aligned}$$

This concludes the proof. ■

Appendix C Proof of Corollary 1

This corollary directly follows from the social welfare with the institutional incentive, defined as the average payoff of the population with an additional deduction of the implementation costs [4]. The social welfare for the state x and the policy u is defined as

$$\begin{aligned} \tilde{\pi}_{\text{all}}(x,u) &= x\tilde{\pi}_C(x,u) + (1-x)\tilde{\pi}_D(x,u) - \alpha x u \\ &= x(b - c - 2\alpha u), \end{aligned}$$

where $\alpha u x$ denotes the costs incurred by the institution. Then, the difference between the social welfare at the target state x_1 and that atthe initial state x_0 under the optimal policy u^* is given by

$$\tilde{\pi}_{\text{all}}(x_1,u^*) - \tilde{\pi}_{\text{all}}(x_0,u^*) = (x_1 - x_0)(b - c - 2\alpha u^*). \tag{C1}$$

An increase in the level of cooperation (i.e. $x_1 > x_0$), does not impair the social welfare $\big($i.e. $\tilde{\pi}_{\text{all}}(x_1,u^*) \geqslant \tilde{\pi}_{\text{all}}(x_0,u^*)\big)$ when the coefficient in (C1) is non-negative, which can be expressed as

$$b - c - 2\alpha u^* \geqslant 0 \iff \alpha u^* \leqslant \frac{b-c}{2}, \tag{C2}$$

where $u^* = \min\big\{\frac{2c}{1-\alpha},\ b\big\}$ is the optimal policy for the case of $x_1 > x_0$. Substituting (B6) into (C2), the condition is equivalent to

$$\begin{cases} \alpha b \leqslant \frac{b-c}{2}, & (b,c,\alpha) \in \big\{(b,c,\alpha) \in \mathbb{R}^3_{\geqslant 0} \big|\ \big(\ b/c > 2,\ \alpha \in [1-2c/b, 1-c/b)\big) \text{ or } \big(b/c \in (1,2],\ \alpha < 1-c/b\big)\big\}, \\ \alpha\frac{2c}{1-\alpha} \leqslant \frac{b-c}{2}, & (b,c,\alpha) \in \big\{(b,c,\alpha) \in \mathbb{R}^3_{\geqslant 0} \big|\ b/c > 2,\ \alpha < 1-2c/b\big\}, \end{cases}$$

which can be classified into three cases.

For $(b,c,\alpha) \in \big\{(b,c,\alpha) \in \mathbb{R}^3_{\geqslant 0} \big|\ b/c > 2,\ \alpha \in [1-2c/b, 1-c/b)\big\}$, the condition (C2) imposes the constraint range on the inefficiency ratio α to be

$$\alpha \in \big[1-2\frac{c}{b}, 1-\frac{c}{b}\big) \cap \big[1-2\frac{c}{b}, \frac{1}{2}-\frac{c}{2b}\big] \implies \alpha \in \big[1-2\frac{c}{b}, \frac{1}{2}-\frac{c}{2b}\big],$$

where the existence of α requires $b/c \leqslant 3$. Therefore, the feasible set for this case can be simplified to

$$\big\{(b,c,\alpha) \in \mathbb{R}^3_+ : b/c \in (2,3],\ \alpha \in \big[1-2\frac{c}{b}, \frac{1}{2}-\frac{c}{2b}\big]\big\}.$$

Similarly, for $(b,c,\alpha) \in \big\{(b,c,\alpha) \in \mathbb{R}^3_+ \big|\ b/c \in (1,2],\ \alpha < 1-c/b\big\}$, the condition (C2) is satisfied if the inefficiency ratio α belongs to

$$\alpha \in \big[0, 1-\frac{c}{b}\big) \cap \big[0, \frac{1}{2}-\frac{c}{2b}\big] \implies \alpha \in \big[0, \frac{1}{2}-\frac{c}{2b}\big],$$

where the existence of such α is obviously guaranteed.

For the last case, the condition (C2) yields that

$$\alpha \in \big[0, 1-2\frac{c}{b}\big) \cap \big[0, \frac{b-c}{b+3c}\big] \implies \begin{cases} \alpha \in \big[0, 1-2\frac{c}{b}\big), & b/c \in (2,3], \\ \alpha \in \big[0, \frac{b-c}{b+3c}\big], & b/c > 3. \end{cases}$$

In summary, the condition under which increasing cooperation does not worsen social welfare is

$$\big\{(b,c,\alpha) \in \mathbb{R}^3_{\geqslant 0} \big|\ \big(b/c \in (1,3],\ \alpha \in \big[0, \frac{1}{2} - \frac{c}{2b}\big]\big) \text{ or } \big(b/c > 3,\ \alpha \in \big[0, \frac{b-c}{b+3c}\big]\big)\big\},$$

and the proof is finished. ■

Appendix D Simulation results

It is worth noting that the optimal policy u^* is derived from the mean field equation, which is also a deterministic dynamical model describing the aggregate behavior of the infinite population. The applicability of the optimal policy for finite but large populations needs to be verified through Monte Carlo simulations.

For the process $\{x^n(\tau)\}_{\tau\in\mathbb{N}}$ of the n-player population under the incentive policy $\{u^n(\tau)|u^n(\tau)=u^*\}_{\tau\in\mathbb{N}}$, we denote the state $x^n(\tau)$ after time scaling as $x^n(\tau\omega/n)$, by abuse of notation. The corresponding cumulative cost for the rescaled state $x^n(\tau\omega/n)$ of the n-player population is denoted by J^n. This time rescaling allows us to compare both the state trajectory $x^n(\cdot)$ against $x^*(\cdot)$, and the cumulative cost J^n against J^*, where $x^*(\cdot)$, J^* are theoretical quantities derived from the mean field equation.

In the following, we will compare $x^*(t)$, J^* with simulation results $\mathbb{E}\big(x^n(\tau\omega/n)\big)$, $\mathbb{E}(J^n)$ for three distinct parameter configurations (b,c,α). Specifically, these configurations correspondingly determine the different values of u^*. The randomness in these variables $x^n(\cdot)$, J^n originates from stochastic fluctuations induced by the finite population size. By comparing the results in Figure D1 and Figure D2, we find that the theoretical approximations can effectively predict the associated simulated results across all three distinct parameter configurations. Against this backdrop, the explanations presented in the subsequent analysis are fully applicable to each of the three configurations.

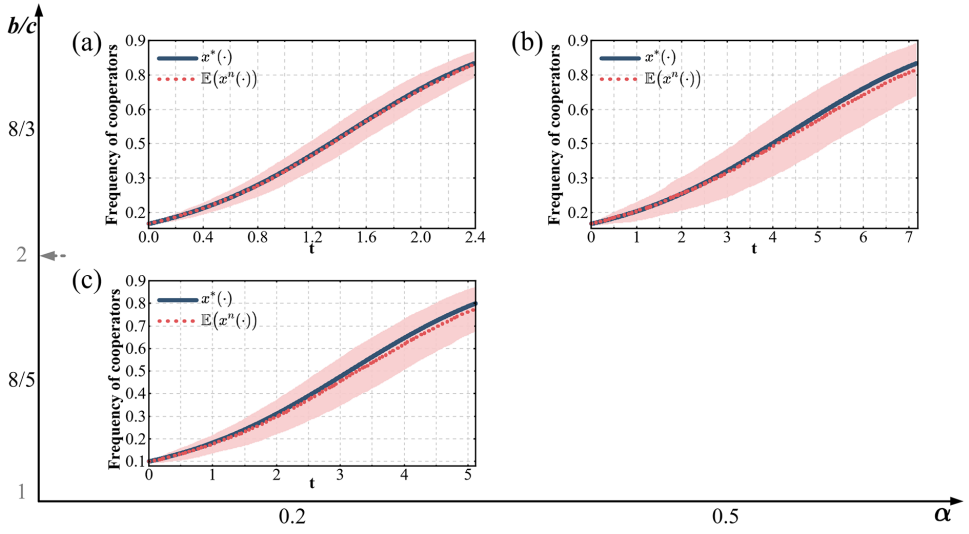

Figure D1 Dynamics of the population state along the time axis t $(=\omega\tau/n)$. All results are derived through 200 independent Monte Carlo simulations initialized with x_0 until the Monte Carlo step τ reaches T^*n/ω. The red lines indicate $\mathbb{E}\big(x^n(\cdot)\big)$ by averaging 200 trials. The shaded regions represent the fluctuation range of the simulation results. The blue lines depict theoretical values $x^*(\cdot)$. Parameters are fixed to $n=10^4$, $\omega=0.01$, $t_0=0$, $x_0=0.1$, $x_1=0.8$. For panel (a), $b=8$, $c=3$, $\alpha=0.2$; for panel (b), $b=8$, $c=3$, $\alpha=0.5$; and for panel (c) $b=8$, $c=5$, $\alpha=0.2$.

As shown in Figure D1, the policy u^* drives the aggregate cooperative behavior $\{\mathbb{E}\big(x^n(\tau\omega/n)\big)\}_{\tau\in\mathbb{N}\cap[0,T^*n/\omega]}$ of the n-player population to increase, which in turn verifies the feasibility of the tax-subsidy incentive with respect to promoting cooperation. The range of random fluctuations reflected by the shaded region exhibits a continuous tendency to grow over time, which stems from the existence of the Wiener process. More precisely, the variance property of the Wiener process grows over time. Despite inherent fluctuations, the theoretical prediction $x^*(t)$ and the simulated average $\mathbb{E}\big(x^n(t)\big)$ exhibit a high degree of consistency. Moreover, the aggregate behavior $\mathbb{E}\big(x^n(t)\big)$ successfully approaches the target state x_1 with high precision at the terminal time T^*. This result indirectly verifies the accuracy of T^*, which is a theoretical quantity derived from (B7). The close match between theoretical and simulated values not only validates the accuracy of the theoretical prediction $x^*(t)$ but also demonstrates the applicability of the optimal policy u^* in large but finite populations. Figure D2 presents a comparison of the other core indicator for the same three parameter configurations. The combination of close

alignment between the theoretical predictions J^* and simulated values $\mathbb{E}(J^n)$ across all configurations, along with negligible error bars, further illustrates the accuracy of the theoretical predictions.

Above all, these results collectively confirm that the theoretical values effectively predict the simulated outcomes in large but finite populations under policy u^*, and this in turn further validates its effectiveness.

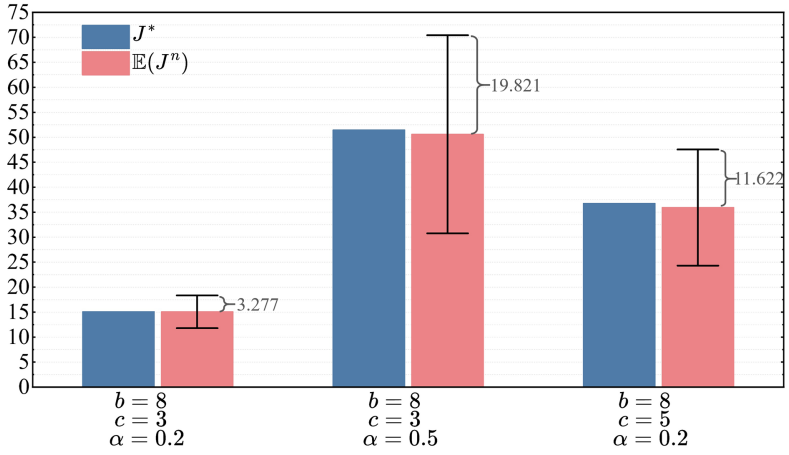

Figure D2 Costs of optimal policies u^* derived from three parameter configurations (b, c, α). The blue bars represent costs J^*, the red bars denote simulated averages $\mathbb{E}(J^n)$, and the error bars indicate the standard deviation across 200 trials. Parameters are fixed to $n = 10^4,\ \omega = 0.01,\ t_0 = 0,\ x_0 = 0.1,\ x_1 = 0.8$.

References

1 Traulsen A, Claussen J C, Hauert C. Coevolutionary dynamics: From finite to infinite populations. Phys Rev Lett, 2005, 95: 238701
2 Gardiner C W. Handbook of Stochastic Methods. Springer Berlin Heidelberg, 1983
3 Liberzon D. Calculus of variations and optimal control theory: a concise introduction. Princeton university press, 2011
4 Han T A, Duong M H, Perc M. Evolutionary mechanisms that promote cooperation may not promote social welfare. J Roy Soc Interface, 2024, 21(220): 20240547