

Research Article

Optimal Tag-Based Cooperation Control for the “Prisoner’s Dilemma”

Rui Dong ^{1,2}, Xinghong Jia ³, Xianjia Wang ^{2,4} and Yonggang Chen ¹

¹School of Mathematics Sciences, Henan Institute of Science and Technology, Xinxiang 453003, China

²Institute of Systems Engineering, Wuhan University, Wuhan 430072, China

³School of E-commerce and Logistics Management, Henan University of Economics and Law, Zhengzhou 450046, China

⁴School of Economics and Management, Wuhan University, Wuhan 430072, China

Correspondence should be addressed to Rui Dong; drui@163.com and Xinghong Jia; xhjia@126.com

Received 5 June 2020; Revised 6 October 2020; Accepted 19 October 2020; Published 11 December 2020

Academic Editor: Danilo Comminiello

Copyright © 2020 Rui Dong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A long-standing problem in biology, economics, and social sciences is to understand the conditions required for the emergence and maintenance of cooperation in evolving populations. This paper investigates how to promote the evolution of cooperation in the Prisoner’s Dilemma game (PDG). Differing from previous approaches, we not only propose a tag-based control (TBC) mechanism but also look at how the evolution of cooperation by TBC can be successfully promoted. The effect of TBC on the evolutionary process of cooperation shows that it can both reduce the payoff of defectors and inhibit defection; although when the cooperation rate is high, TBC will also reduce the payoff of cooperators unless the identified rate of the TBC is large enough. An optimal timing control (OTC) of switched replicator dynamics is designed to consider the control costs, the cooperation rate at terminal time, and the cooperator’s payoff. The results show that the switching control (SC) between an optimal identified rate control of the TBC and no TBC can properly not only maintain a high cooperation rate but also greatly enhance the payoff of the cooperators. Our results provide valuable insights for some clusters, for example, logistics parks and government, to regard the decision to promote cooperation.

1. Introduction

Animal Dispersion in Relation to Social Behavior was published by Wynne-Edwards and looked specifically at the “Evolution of cooperation” [1]. Why should an individual help another person who is a potential competitor in the struggle for survival? This question is listed in “Science” Magazine as one of the 25 core problems of the 125 scientific challenges proposed by scientists from all over the world.

The evolution of cooperation is an enduring conundrum in biology, mathematics, and social sciences. Natural selection opposes cooperation unless some mechanisms are at work to promote the evolution of cooperation. The Prisoner’s Dilemma game (PDG), which represents an extreme case, has emerged as one of the most promising mathematical areas in the study of cooperation.

In the PDG, two players simultaneously decide whether to cooperate (C) or to defect (D). If one player cooperates,

the other player can choose between cooperation which yields R (the reward for mutual cooperation) or defection which yields T (the temptation to defect). On the other hand, if one player defects, the other player can choose between cooperation which yields S (the sucker’s payoff) or defection which yields P (the punishment for mutual defection), where $T > R$ and $P > S$. The donor-recipient game (DRG) is a special case of the PDG. In the DRG, a cooperator is someone who pays a cost, c , for another individual to receive a benefit, b . A defector has no cost and does not deal out benefits. By comparing the PDG and DRG, we obtain $R = b - c$, $T = b$, $S = -c$, and $P = 0$. If the game is only played once, then each player gets a higher payoff from D than from C , regardless of what the other player does. So, natural selection implies competition and, therefore, opposes cooperation unless a specific mechanism is at work.

Based on Hamilton’s rule [2], natural selection can favor cooperation if the donor and the recipient of an altruistic act

are genetic relatives [3–7]. The donor may preferentially donate a benefit to the recipient through “kin recognition”. Explanations of cooperation between nonkin include “direct reciprocity” [8–13], “indirect reciprocity” [14–16], “network reciprocity” [17], “group selection” [18], “optional participation mechanism” [19, 20], and “punishment mechanism” [21–23].

Direct reciprocity was proposed by Trivers [8] and developed as a mechanism for the evolution of cooperation [9, 24]. In the repeated Prisoner’s Dilemma, the same two individuals repeatedly encounter each other for some rounds. If one cooperates now, the other may cooperate later. Hence, he or she *might* cooperate [10]. Axelrod and Hamilton proposed a model of the evolution of cooperation based on the iterated Prisoner’s Dilemma [11]. In two computer tournaments, Axelrod discovered that the “winning strategy” was the simplest of all, tit-for-tat (TFT). TFT is a program where one player uses *C* on the first move of the game and, then, plays whatever the other player chose in the previous move. This simple concept captured the fascination of many enthusiasts of the repeated Prisoner’s Dilemma, and a number of empirical and theoretical studies were inspired by Axelrod’s groundbreaking work [12, 13, 25]. The presence of the decoy increased willingness of volunteers to cooperate in the first step of each game, leading to subsequent propagation of such willingness by (noisy) tit-for-tat [25]. Reference [26] pointed that resilient cooperators could sustain cooperation indefinitely in the finitely repeated Prisoner’s Dilemma. The Prisoner’s Dilemma experiment showed that the termination rules and the (expected) length of the game significantly increased cooperation in [27].

Indirect reciprocity does not rely on repeated encounters between the same two individuals. An individual can establish a good reputation by helping someone which can increase the chance of receiving help from others. Indirect reciprocity has substantial cognitive demands. For example, language is needed to gain the information and spread the gossip associated with indirect reciprocity so that only humans seem to engage in the full complexity of the game. Differing from indirect reciprocity based on reputation, kin can be recognized through familiarity based on environmental or learned cues or through pattern matching based on some inherited trait. If individuals display a heritable marker or a tag, they preferentially cooperate with partners who share their own marker. So, tag-based donation does not have the substantial cognitive demands that a reputation mechanism based on indirect reciprocity has. Cooperation can also arise when individuals donate to others who are sufficiently similar to themselves in some arbitrary characteristic. Such a characteristic, or “tag,” for example, a green beard [28, 29], can be observed by others. Hamilton illustrated how this mechanism worked even when individuals were not genealogical kin [2]. This is known as the green beard effect. For tag-based donations, it is not necessarily required to remember past encounters. In the recent years, the green beard effect has become a favorite topic among both evolutionary biologists and sociologists [30–36]. For example, Tian et al. and Tian et al. [30, 31] proposed Vcash as

a reputation framework for identifying denial of traffic service to resolve the trustworthiness problem, where the result of verified traffic event notification acted as a “tag”. Taylor and Nowak proposed nonuniform interaction rates (interaction rates dependent on the strategies) that allowed the coexistence of cooperators and defectors in the Prisoner’s Dilemma [37]. The analytical models of evolution of cooperation were given when nonuniform interaction rates were introduced [38]. Both “kin recognition” and “tag-based donation” lead to nonuniform interaction rates.

By combining evolutionary processes with differential equations, the links between stability and optimization have been researched and a mathematical framework has been built [39, 40]. In multiagent systems, an individual is regarded as an agent who autonomously regulates his behavior according to his benefit. Agents play games with their neighbors through local interaction. The evolution of cooperation is an adaptive coordinated control process which is a controllable, intelligent, and autonomous decision process. From the perspective of evolutionary games, the following objectives, *inter alia*, have been achieved through the designing of the control law: optimizing individual cost function [39], designing consensus control of stochastic multiagent systems [41], and promoting the evolution of cooperation in social dilemmas [23, 38].

Motivated by the abovementioned discussion, through combining direct reciprocity, indirect reciprocity, tag-based donation, and optimal control theory, a tag-based control (TBC) is proposed. The goal of this paper is to study the feasibility of promoting cooperation by TBC and designing an appropriate identified rate for operators. As part of this, we discuss the different effectiveness and drawbacks between by identifying cooperation and by identifying defection for promoting the evolution of cooperation. Using the optimal control theory, we examine whether an operator should design TBC mechanisms for the evolution of cooperation or give up TBC temporarily for increasing the group’s payoff. We show that an operator can design an optimal timing control (OTC) to optimize the control costs, the cooperation rate, and the cooperator’s payoff.

The main contributions of this paper are as follows. (1) For the first time, a TBC that promotes the evolution of cooperation in the Prisoner’s Dilemma is proposed. (2) Advantages and disadvantages of identifying cooperation and identifying defection in TBC are investigated. (3) An OTC is designed to better promote cooperation rates and greatly enhance the group’s payoff. (4) From a research standpoint, this work contributes to evolutionary game theory and optimal control theory.

The paper is organized as follows. In Section 2, we obtain the payoff matrix and the replicator dynamics of the Prisoner’s Dilemma with TBC. By applying the results of Section 2, we discuss the problem of evolution of cooperation in the Prisoner’s Dilemma with TBC in Section 3. In the next two sections, the optimal TBC is designed. We design the optimal identified rate of promoting cooperation in Section 4. The effectiveness of TBC is illustrated and an OTC is given in Section 5. We provide some concluding remarks and discussion in the last section.

2. The Replicator Dynamics of the Prisoner's Dilemma with TBC

In the Prisoner's Dilemma game, two individuals can each either cooperate (C) or defect (D). The payoff matrix of the Prisoner's Dilemma is as follows:

$$\begin{array}{cc} & \begin{array}{cc} C & D \end{array} \\ \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R & S \\ T & P \end{pmatrix}, \end{array} \quad (1)$$

where

$$T > R > P > S, \quad (2)$$

$$2R > T + S. \quad (3)$$

No matter what the other does, the selfish choice of defection yields a higher payoff than cooperation, so a player will choose defection no matter whether his opponent chooses cooperation or defection. However, if both defect, both get P rather than the larger value of R that they both could have gotten if they had both cooperated. In other words, if both defect, both do worse than if both had cooperated. Hence, the game poses a dilemma. In repeated Prisoner's Dilemma games, if constraint (3) cannot be satisfied, the social efficiency that all players always choose cooperation is less than that they agree to choose cooperation and defection in turn.

Let $\Omega = \{1, 2, \dots, N\}$ be a group of game participants, where N is a sufficiently large natural number. Consider the symmetric static games of complete information between two individuals. The pure strategy set for each individual in group Ω is denoted by $\Lambda' = \{C, D\}$. For notational convenience, we will label every individual's pure strategies by positive integers. Hence, the pure-strategy set of each individual is written as $\Lambda = \{1, 2\}$. A vector of pure-strategy profile is denoted as $s = \{s_1, s_2\}$, where s_i is a pure strategy for individual i . The pure strategy space is $\Lambda \times \Lambda$, where $\Lambda \times \Lambda$ is the Cartesian product of the pure strategy set Λ . A probability distribution (x, y) over the pure-strategy set Λ of an individual i is defined as a hybrid strategy for the individual i , where x and y are the probabilities assigned to the individual's pure strategy C and D , respectively. We denote $z = (x, y)$. The hybrid strategy space of each individual is the simplex $\Delta = \{z \in R_+^2: x + y = 1\}$, where R_+^2 is a subset of the two-dimensional vectors with all elements being positive. At any point of time t , let $x(t)$ and $y(t)$ be the rates at which individuals choose strategy C and D in group Ω , respectively. Then, the corresponding group state is $z(t) = (x(t), y(t))$, where $x(t), y(t) \in (0, 1)$ and $x(t) + y(t) = 1$. Therefore, the group state is formally equivalent to the hybrid strategy. Because the state $z(t)$ is completely determined by its component $x(t)$ or $y(t)$, this paper discusses the variable $x(t)$ only.

In the group Ω , the individuals who choose strategy D have a higher average fitness than the others. Therefore,

selection acts to increase the relative abundance of strategy D . After sometime, cooperation will vanish from the group. So, strategy D is the only evolutionarily stable strategy (ESS) unless a specific mechanism is at work.

2.1. The Game under Labeling. This paper studies the special mechanisms for the evolution of cooperation in some clusters in the economic society when faced with the Prisoner's Dilemma, for example, in logistics parks. For such a large number of individuals from different sources, it is difficult to build mutual trust because of the high cost of fully understanding and accurately knowing each other's willingness to cooperate. In this case, the mechanisms for promoting the evolution of cooperation, such as direct reciprocity, indirect reciprocity, punishment mechanism, and nonuniform interaction, are difficult to meet. A third party in the cluster, though, can play the role of collecting information and rewarding or punishing enterprise behavior to promote mutual recognition efficiently. Such examples might be park managers, platform promoters, park management committees, industry associations, and so on.

In view of this, we assume that there exists a controller (an operator) outside of group Ω who can identify defection and cooperation. The controller labels "cooperator" or "defector" for everyone. The tags labeled by the controller affect the strategy choice of individuals in the following games until the next identification.

Definition 1. A person or an organization is called a controller if he/she can affect the choice of strategy of individuals through identifying the actions of individuals in past games and reveal the results of identification in group Ω .

The identification implemented by the controller can be regarded as tagging individuals. Strategy choices of all individuals in group Ω are decided by their tags in the following games.

Definition 2. One round game is defined as a game set which consists of a game identified by a controller and all games tagged by this identification.

Assume the number of rounds is infinite. In one round, every individual plays $m + 1$ period games, where m is a positive integer. The first period is identified by the controller in one round. All individuals are tagged before the second period game begins. In the following m periods, all individuals know their tags and can distinguish tags of others.

Assumption 1. The individuals are bounded rationality and any individual encounters other individuals randomly with equal probability.

This paper assumes that the individuals (enterprises or persons) are in certain clusters (such as logistics parks) or bilateral platforms in the economic society. The indirect reciprocity has substantial cognitive demands for players. The tag mechanism for the evolution of cooperation demands that individuals own and display a heritable marker or a tag. Due to the large number of individuals in the

cluster, the interaction between them is random, and it is difficult to form long-term repeated games. Therefore, it is impossible to establish a direct reciprocity mechanism based on repeated games. In the early stage of the formation of clusters or bilateral platforms, due to the lack of mutual understanding and trust relationship between individuals, the cost of collecting information and displaying individual reputation is very high, so indirect reciprocity and tag mechanism cannot work unless a person or organization helps the players. We define this person or organization as the controller (Definition 1). The identification implemented by the controller can be regarded as tagging individuals. In the first period of a round game, the controller identifies individuals and displays the results.

Based on Assumption 1, for each individual, the probability of interacting with a cooperator or a defector is $x(t)$ or $1 - x(t)$, respectively.

Definition 3. Define an individual as a cooperator if he/she chooses cooperation without control.

At time t , the rate of cooperators in the group Ω is $x(t)$. So, the state $x(t)$ is defined as the group *cooperation rate*. The cooperator defined by Definition 3 can be equated with the individual who always chooses cooperation without control based on Assumption 1. Obviously, an individual is a defector if he is not a cooperator under Definition 3.

Assumption 2. Assume that the identification service is not perfect. The probabilities that the cooperators and defectors are correctly identified by the controller are β_1 and β_2 , respectively, where $\beta_i \in (0, 1)$, $i = 1, 2$.

Based on Assumption 2, β_1 is called the cooperation identified rate and β_2 is called the defection identified rate. We denote the tags of ‘‘cooperator’’ and ‘‘defector’’ as ‘‘ C' ’’ and ‘‘ D' ’’. In one round, Ω is divided into 4 parts noted by Ω_i ($i = 1, 2, 3, 4$). Ω_1 and Ω_2 are sets of cooperators and defectors labeled C' , respectively. Analogously, Ω_3 and Ω_4 are sets of cooperators and defectors labeled as D' , respectively. Furthermore, $\cup_{i=1}^4 \Omega_i = \Omega$, $\Omega_i \cap \Omega_j = \Phi$, where Φ is an empty set. By denoting individual i labeled j' as $j' | i$ ($i, j = C, D$), one can know that $C' | C \in \Omega_1$, $C' | D \in \Omega_2$, $D' | D \in \Omega_3$, and $D' | C \in \Omega_4$. According to Assumption 2, $p(C' | C) = \beta_1$, $p(D' | C) = 1 - \beta_1$, $p(D' | D) = \beta_2$, and $p(C' | D) = 1 - \beta_2$, where $p\{\cdot\}$ indicates the probability of event $\{\cdot\}$ happening.

Assumption 3. In every round game, cooperators choose C and defectors choose D in the first period. In the following m periods, the cooperators labeled D' and all defectors always choose D . Cooperators labeled C' choose C if he/she encounters an individual labeled C' and chooses D if he/she encounters an individual labeled D' .

Definition 4. The control of a controller dominating individuals' strategy choice through tagging is called TBC. The individual identified as the cooperator is tabled as C' , and the individual identified as the defector is tabled as D' by the controller. The TBC rule is designed as follows: the cooperators labeled D' and all defectors always choose D .

Cooperators labeled C' choose C if he/she encounters an individual labeled C' and chooses D if he/she encounters an individual labeled D' .

Based on Assumption 3, the controller can dominate an individual's strategy choice through tagging individuals. At time t , the cooperation rate is $x(t)$ in group Ω . In the first period, the ratio of C to total strategies is equal to the cooperation rate $x(t)$. In the following m periods, an individual's strategy is decided by tags according to Assumption 3. The ratio of C to total strategies is not the cooperation rate $x(t)$ but $\beta_1 x [\beta_1 x + (1 - \beta_2)(1 - x)]$. One can know that $\beta_1 x [\beta_1 x + (1 - \beta_2)(1 - x)] = \beta_1 x [1 - x(1 - \beta_1) - \beta_2(1 - x)] < x$.

Assumption 3 is reasonable by Assumption 1 and Definition 1. From the abovementioned discussion, we know that C not only denotes cooperation but also denotes a cooperator and D not only denotes defection but also denotes a defector. Numbers of individuals in $\Omega_1, \Omega_2, \Omega_3$, and Ω_4 are $\beta_1 x N$, $(1 - \beta_2)(1 - x)N$, $\beta_2(1 - x)N$, and $(1 - \beta_1)xN$, respectively. Based on Assumption 2 and 3, in the first period, cooperators choose C and defectors choose D . In the following m periods, two individuals choose C , if and only if they are in Ω_1 . The strategy pair (C, C) is contributed. When one player from Ω_1 encounters a player from Ω_2 , a cooperator chooses C and a defector chooses D . In addition to the abovementioned two cases, the strategy pair (D, D) will be formed. According to Taylor and Nowak [31] and Dong et al. [32], the probabilities of (C, C) , (D, D) , and (C, D) are $\beta_1^2 x^2$, $1 - \beta_1^2 x^2 - 2\beta_1(1 - \beta_2)x(1 - x)$, and $\beta_1(1 - \beta_2)x(1 - x)$ with TBC. The larger β_1 and β_2 means the higher probability of (C, C) and the lower probability (C, D) . So, the larger β_1 and β_2 are in accordance with the better effect of promoting the evolution of cooperation. But, under uniform interaction rates, the corresponding probabilities are x^2 and $(1 - x)^2$, respectively, so an individual's encounter constrained by TBC no longer abides by uniform interaction rates.

Based on Assumptions 2 and 3 and Definition 4, cooperation identified rate β_1 and defection identified rate β_2 reflect the accuracy of TBC. We denote β' as the probability of another encounter between the same two individuals in the repeated Prisoner's Dilemma and β'' as the probability of knowing someone's reputation in indirect reciprocity. According to [31], the condition that direct reciprocity (indirect reciprocity) can lead to the evolution of cooperation depends on the parameters of payoff matrix (1) and β' (β''). But, by Assumption 1, β' is very small because the group size, N , is large enough. On the other hand, the high cost of fully understanding and accurately knowing each other's willingness to cooperate leads to that β'' is very small also. So, direct reciprocity and indirect reciprocity cannot promote cooperation in the setting that this paper considers. In every round game, cooperation identified rate β_1 and defection identified rate β_2 can act as β' of direct reciprocity or β'' of indirect reciprocity.

Assumption 4. The evolution of group state x is carried out according to the replicator. An individual replication occurs after each round game ends and before the next round begins. An individual's payoff refers to the average total payoff in this round game.

Assumption 4 means the offspring of cooperators are still cooperators and the offspring of defectors are still defectors. So, the offsprings of C or D are disconnected with their parent's tag. By Assumption 3, defectors always choose D , but the cooperator does not always choose C .

"Tag" does a mapping σ between Ω and $\{C', D'\}$, i.e., $\sigma: \Omega \rightarrow \{C', D'\}$. By Assumption 2, $p\{\sigma(C) = C'\} = \beta_1$, $p\{\sigma(C) = D'\} = 1 - \beta_1$, $p\{\sigma(D) = D'\} = \beta_2$, and $p\{\sigma(D) = C'\} = 1 - \beta_2$.

By calculation, the payoff of cooperator (defector) in the p th period of the r th round game is independent of r but determined by the cooperation rate x . Let $v_{pth,C}(x)$ ($v_{pth,D}(x)$) denote the payoff of a cooperator (defector) in the p th period. Let $w_C(x)$ ($w_D(x)$) denote total payoff of the cooperator (defector) in one round game; we get

$$\begin{aligned} w_C(x) &= v_{1th,C}(x) + \sum_{p=2}^{m+1} v_{pth,C}, \\ w_D(x) &= v_{1th,D}(x) + \sum_{p=2}^{m+1} v_{pth,D}. \end{aligned} \quad (4)$$

Let $v_{pth,j'/C}(x)$ ($v_{pth,j'/D}(x)$) denote the payoffs of the cooperator (defector) labeled j' ($j' = C, D$) in the p th period game; we get

$$\begin{aligned} v_{pth,C}(x) &= p(\sigma(C) = C')v_{pth,C'|C} + p(\sigma(C) = D')v_{pth,D'|C}, \\ v_{pth,D}(x) &= p(\sigma(D) = C')v_{pth,C'|D} + p(\sigma(D) = D')v_{pth,D'|D}, \\ & p = 2, 3, \dots, m+1. \end{aligned} \quad (5)$$

By Assumptions 1 and 2, the probability of anyone encountering a cooperator labeled C' is $\beta_1 x$ and the

probability of anyone encountering a cooperator labeled D' is $(1 - \beta_1)x$; the probability of anyone encountering a defector labeled C' or D' is $(1 - \beta_2)(1 - x)$ or $\beta_2(1 - x)$. One can get

$$\begin{aligned} v_{1th,C}(x) &= Rx + S(1 - x), \\ v_{pth,C'/C}(x) &= \beta_1 xR + [(1 - \beta_1)x + \beta_2(1 - x)]P \\ & \quad + (1 - \beta_2)(1 - x)S, \\ v_{pth,D'/C}(x) &= P, \quad p = 2, 3, \dots, m+1, \end{aligned} \quad (6)$$

$$\begin{aligned} v_{1th,D}(x) &= xT + (1 - x)P, \\ v_{pth,C'/D}(x) &= \beta_1 xT + (1 - \beta_1)xP, \\ v_{pth,D'/D}(x) &= P, \quad p = 2, 3, \dots, m+1. \end{aligned} \quad (7)$$

Together, (4)–(7) determine $w_C(x)$ and $w_D(x)$ as follows:

$$\begin{aligned} w_C(x) &= (1 + m\beta_1^2)xR \\ & \quad + [\beta_1(1 - \beta_1 - \beta_2)x + 1 - \beta_1(1 - \beta_2)]mP \\ & \quad + \frac{n!}{r!(n-r)!} [1 + m\beta_1(1 - \beta_2)](1 - x)S, \\ w_D(x) &= v_{1,D}(x) \\ & \quad + m[\beta_2 v_{pth,D'/D}(x) + (1 - \beta_2)v_{pth,C'/D}(x)], \\ & = [1 + m\beta_1(1 - \beta_2)]xT \\ & \quad + \{1 - x + m[1 - \beta_1(1 - \beta_2)x]\}P. \end{aligned} \quad (8)$$

By Assumption 3, an individual's average payoff in the r th round game is $xw_C(x) + (1 - x)w_D(x)$.

Summing up the abovementioned discussion, the total payoff matrix (1) of a round game under TBC is

$$\begin{bmatrix} (1 + m\beta_1^2)R + (1 - \beta_1^2)mP & [1 + m\beta_1(1 - \beta_2)]S + [1 - \beta_1(1 - \beta_2)]mP \\ [1 + m\beta_1(1 - \beta_2)]T + [1 + \beta_1(1 - \beta_2)]mP & mP \end{bmatrix}. \quad (9)$$

In fact, in the latter m periods, the hybrid strategy and the group state are not equivalent because a cooperator constrained by Assumption 3 does not always choose C . So, the payoff a_{ij} in payoff matrix (9) does not indicate the payoff of individual i encountering strategy j , where $i, j = C, D$. However, we still use payoff matrix (9) for the following two reasons: (1) the payoff $w_i(x)$ of individual Ω_1 is still linear with the group state $x(t)$; (2) a cooperator's average total payoff under TBC is $(1 + m\beta_1^2)R + [1 - \beta_1^2]mP$ if he enters a group of cooperators or is $[1 + m\beta_1(1 - \beta_2)]S + [1 - \beta_1(1 - \beta_2)]mP$ if he enters a group of defectors. Similarly, a defector's average total payoff under TBC is $[1 + m\beta_1(1 - \beta_2)]T + [1 - \beta_1(1 - \beta_2)]mP$ if he enters a group of cooperators or is P if he enters a group of defectors.

If $\beta_1 = 1, \beta_2 = 0$, i.e., all individuals are identified as cooperators, the game with payoff matrix (9) is equivalent to a Prisoner's Dilemma repeated $m+1$ times with payoff matrix (1). By Assumption 3, cooperators always choose C and defectors always choose D .

2.2. The Replicator Dynamics with TBC. At time t , we suppose the number of cooperators is M and the number of defectors is $N - M$. In the next Δt , under Assumption 4, the cooperators reproduce $w_C M \Delta t$ offspring and the defectors reproduce $w_D (N - M) \Delta t$ offspring. The total number of death is $[w_C M + w_D (N - M)] \Delta t$. The number of cooperators death is $[w_C M + w_D (N - M)] x \Delta t$, and the number of defectors death is $[w_C M + w_D (N - M)] (1 - x) \Delta t$. Then, $\Delta x = \{w_C M \Delta t - [w_C M + w_D (N - M)] x \Delta t\} / N = (w_C - w_D) x (1 - x) \Delta t$. Then, the replicator dynamics under TBC is as follows:

$$\begin{aligned} \dot{x}(t) &= [w_C(x) - w_D(x)](1 - x)x, \\ & = [(R - S - T + P)x + S - P](1 - x)x \\ & \quad + m\{\beta_1 R + (2 - \beta_1 - 2\beta_2)P - (1 - \beta_2)(S + T)\}x \\ & \quad - (1 - \beta_2)(P - S)\beta_1(1 - x)x. \end{aligned} \quad (10)$$

Assumption 5. The Prisoner's Dilemma tagged repeats enough periods, i.e., $m \gg 1$.

By Assumption 5, the total payoff matrix (9) in a round game approximately equals the following matrix:

$$\begin{bmatrix} m\beta_1^2 R + (1 - \beta_1^2)mP & [1 + \beta_1(1 - \beta_2)]mp + m\beta_1(1 - \beta_2)S \\ [1 + \beta_1(1 - \beta_2)]mP\beta_1(1 - \beta_2)mT & mP \end{bmatrix}. \quad (11)$$

The replicator dynamics (10) approximates the following dynamics:

$$\begin{aligned} \dot{x}(t) = m\{ & [\beta_1 R + (2 - \beta_1 - 2\beta_2)P - (1 - \beta_2)(S + T)]x \\ & - (1 - \beta_2)(P - S)\}\beta_1(1 - x). \end{aligned} \quad (12)$$

Definition 5 (See [42]). A state $(\bar{x}, \bar{y}) \in \Delta$ is an ESS if for all other states $(x, y) \in \Delta$ either (i) $u[(x, y), (\bar{x}, \bar{y})] < u[(\bar{x}, \bar{y}), (\bar{x}, \bar{y})]$ or (ii) $u[(x, y), (\bar{x}, \bar{y})] = u[(\bar{x}, \bar{y}), (\bar{x}, \bar{y})]$ and $u[(x, y), (x, y)] < u[(\bar{x}, \bar{y}), (x, y)]$, where $u(\cdot_1, \cdot_2)$ denotes the individual's payoff with strategy \cdot_1 under the strategy pair (\cdot_1, \cdot_2) .

Definition 6 (See [43]). A fixed point $(\bar{x}, \bar{y}) \in \Delta$ in replicator dynamics is an evolutionary equilibrium (EE) if it is locally asymptotically stable, i.e., every open neighborhood $K \subset \Delta$ of the point (\bar{x}, \bar{y}) has the property that every path starting sufficiently close to (\bar{x}, \bar{y}) remains in K and converges asymptotically to (\bar{x}, \bar{y}) .

Definition 7 (See [43]). The largest open set of points whose evolutionary paths converge to a given EE is called its basin of attraction.

In order to simplify the calculation, $P = 0$ is assumed in the following discussion. Doing so would result in a simpler model, without significant qualitative or directional changes to our key results about evolution of cooperation. Obviously, the parameters T, R , and S satisfy $T > R > 0 > S$. Based on (11), the payoff matrix for a period under TBC is transformed as follows:

$$\begin{bmatrix} C & D \\ C & \begin{pmatrix} \beta_1^2 R & \beta_1(1 - \beta_2)S \end{pmatrix} \\ D & \begin{pmatrix} \beta_1(1 - \beta_2)T & 0 \end{pmatrix} \end{bmatrix}. \quad (13)$$

Regardless of time scale, m does not affect either ESS or EE and does not even affect the evolutionary path of a replicator's dynamics, so we can rewrite the replicator dynamics (12) as follows:

$$\dot{x}(t) = \{[\beta_1 R - (1 - \beta_2)(S + T)]x + (1 - \beta_2)S\}\beta_1(1 - x). \quad (14)$$

Two specific instances are given in the following.

Case 1. Only the defectors are identified, and let $p(D'/D) = \beta_2$. Ω is divided into two parts. One is the

set of defectors tagged D' . The other is the set of individuals without tag. The corresponding part of Assumption 3 is modified as follows: a cooperators chooses C if he encounters an individual without tag but chooses D if he encounters an individual tagged D' . Defectors always choose defect. In this case, the payoff matrix is as follows:

$$\begin{bmatrix} C & D \\ C & \begin{pmatrix} R & (1 - \beta_2)S \end{pmatrix} \\ D & \begin{pmatrix} (1 - \beta_2)T & 0 \end{pmatrix} \end{bmatrix}. \quad (15)$$

Case 2. Only the cooperators are identified, and let $p(C'/C) = \beta_1$. Ω is divided into two parts. One is the set of cooperators tagged C' . The other is the set of individuals without a tag. The corresponding part of Assumption 3 is modified as follows: the cooperators chooses C if he encounters an individual with tag C' and chooses D if he encounters an individual without a tag. Defectors always choose defect. In this case, we obtain the payoff matrix as follows:

$$\begin{bmatrix} C & D \\ C & \begin{pmatrix} \beta_1^2 R & 0 \end{pmatrix} \\ D & \begin{pmatrix} 0 & 0 \end{pmatrix} \end{bmatrix}. \quad (16)$$

By comparing matrices (13) and (15), we see that Case 1 is equivalent to cooperation identified rate $\beta_1 = 1$ in matrix (13), i.e., the controller accurately identifies cooperators. It follows that Case 1 is in accordance with control based on identifying defection. Correspondingly, Case 2 means the controller identifies defectors accurately.

From the operator's perspective, the relevant decisions are (1) whether to design TBC; (2) identifying which individual is more effective, cooperators or defectors; and (3) how to design cooperation identified rate β_1 and defection identified rate β_2 .

3. Evolution of Cooperation for the Prisoner's Dilemma with TBC

3.1. Short-Term Effects Caused by TBC. Based on Assumption 3, an individual chooses a strategy according to a tag. Can a tag improve the correctness of the choice?

Proposition 1. Let i/j' denote the individual i labeled j' , where $i, j = C, D$. We denote $p(C/C')$ as the probability that the individual tagged C' is a cooperator and $p(D/D')$ as the probability that the individual tagged D' is a defector. Then, $p(C/C') > p(C)$ and $p(D/D') > p(D)$, if and only if $\beta_1 + \beta_2 > 1$.

Proof. According to Assumption 2, $p(C'/C) = \beta_1$ and $p(D'/D) = \beta_2$. Denote the probability of an individual labeled i' as $p(i')$. Obviously, $p(C') = x\beta_1 + (1-x)(1-\beta_2)$ and $p(D') = x(1-\beta_1) + (1-x)\beta_2$. So, one can obtain posterior probability $p(C/C')$ and $p(D/D')$ as follows:

$$\begin{aligned} p\left(\frac{C}{C'}\right) &= \frac{x\beta_1}{x\beta_1 + (1-x)(1-\beta_2)}, \\ p\left(\frac{D}{D'}\right) &= \frac{(1-x)\beta_2}{x(1-\beta_1) + (1-x)\beta_2}. \end{aligned} \quad (17)$$

Thus, $p(C/C') > p(C)$ and $p(D/D') > p(D)$, if and only if $\beta_1 + \beta_2 > 1$.

Now, we discuss an individual's payoff with TBC. Let $f_i(x, \beta_1, \beta_2)$ be the payoff of individual i with TBC and $g_j(x)$ be the payoff of individual j without TBC, where $i, j = C, D$. \square

Proposition 2. If $\beta_1 + \beta_2 > 1$, then $f_C(x, \beta_1, \beta_2) - f_D(x, \beta_1, \beta_2) > g_C(x) - g_D(x)$.

Proof. From matrices (1) and (13), one can obtain

$$\begin{aligned} g_C(x) &= (R-S)x + S, \\ f_C(x, \beta_1, \beta_2) &= [\beta_1^2 R - \beta_1(1-\beta_2)S]x + \beta_1(1-\beta_2)S, \end{aligned} \quad (18)$$

$$\begin{aligned} g_D(x) &= Tx, \\ f_D(x, \beta_1, \beta_2) &= \beta_1(1-\beta_2)Tx. \end{aligned} \quad (19)$$

It follows from (18) and (19) that $f_C(x, \beta_1, \beta_2) - g_C(x)$ and $f_D(x, \beta_1, \beta_2) - g_D(x)$ are all linear to x . Since $f_C(0, \beta_1, \beta_2) - g_C(0) = -S[1 - \beta_1(1 - \beta_2)]$ and $f_D(0, \beta_1, \beta_2) - g_D(0) = 0$, by $-S > 0$, it can be derived that

$$f_C(0, \beta_1, \beta_2) - g_C(0) > f_D(0, \beta_1, \beta_2) - g_D(0). \quad (20)$$

By $f_C(1, \beta_1, \beta_2) - g_C(1) = (\beta_1^2 - 1)R$ and $f_D(1, \beta_1, \beta_2) - g_D(1) = [\beta_1(1 - \beta_2) - 1]T$, if $\beta_1 + \beta_2 > 1$, then $\beta_1(1 - \beta_2) - 1 < \beta_1^2 - 1 < 0$. So, we get

$$f_C(1, \beta_1, \beta_2) - g_C(1) > f_D(1, \beta_1, \beta_2) - g_D(1). \quad (21)$$

In summary, $f_C(x, \beta_1, \beta_2) - g_C(x) > f_D(x, \beta_1, \beta_2) - g_D(x)$ if $\beta_1 + \beta_2 > 1$.

Based on Proposition 1, TBC can help cooperators not only identify cooperators but also reduce the risk of unilateral fraud by defectors if $\beta_1 + \beta_2 > 1$. According to Proposition 2, for any cooperation rate $x \in (0, 1)$, TBC satisfying $\beta_1 + \beta_2 > 1$ can inhibit defection. But, Propositions 1 and 2 do not mean TBC can promote a cooperator's payoff even when $\beta_1 + \beta_2 > 1$. This judgment is confusing. Propositions 3 and 4 will demonstrate. \square

Proposition 3. We denote $*_{ix} = \partial *_{i}/\partial x$, where $* = f, g$; $i = C, D$. With increasing cooperation rate x , some properties of the individual's payoff are as follows:

- (i) $*_{ix} > 0$
- (ii) $f_{ix}(x, \beta_1, \beta_2) - g_{ix}(x) < 0$

Proof. From (18) and (19), according to $T > R > 0 > S$, we get $*_{ix}$ and $f_{Cx}(x, \beta_1, \beta_2) - g_{Cx}(x) = (\beta_1^2 - 1)R + [1 - \beta_1(1 - \beta_2)]S$. Combining with $R > 0$, $S < 0$ and $\beta_1, \beta_2 \in (0, 1)$, we derive $f_{Cx}(x, \beta_1, \beta_2) - g_{Cx}(x) < 0$. Similarly, we can get $f_{Dx}(x, \beta_1, \beta_2) - g_{Dx}(x) < 0$. So, with increases in cooperation rate x , the payoff of all individuals with TBC increases less than their payoff without TBC.

By Proposition 3, the fact that all individual's payoff increases in cooperation rate x has nothing to do with TBC or not. But, the rate of increase of an individual's payoff with TBC is lower than its payoff without TBC. In this sense, TBC reduces group efficiency when the rate of group cooperation rate increases. Why do we design TBC in the Prisoner's Dilemma? In fact, without special mechanisms for promoting the evolution of cooperation, the group cooperation rate will always gradually decrease until cooperators are completely expelled from the group. So, the effect of TBC on the payoff of individuals and the evolution of cooperation is complex. The effect of TBC on the cooperator's payoff and cooperation evolution involves β_1, β_2 , parameters R, T, S , and group cooperation rate x . We denote $x_1(\beta_1, \beta_2) \triangleq (([1 - \beta_1(1 - \beta_2)](-S))/((1 - \beta_1^2)R + [1 - \beta_1(1 - \beta_2)](-S)))$, $x_2(\beta_1, \beta_2) \triangleq (((1 - \beta_2)(-S))/(2\beta_1 R + (1 - \beta_2)(-S)))$. \square

Proposition 4. The effects of TBC on group cooperation rate x and an individual's payoff are as follows:

- (i) For any $x \in (0, 1)$, the defector's payoff with TBC is always lower than its payoff without TBC.
- (ii) A cooperator's payoff with TBC is higher than its payoff without TBC when the group cooperation rate x is lower than $x_1(\beta_1, \beta_2)$ but is lower than its payoff without TBC when x is higher than $x_1(\beta_1, \beta_2)$, i.e., $\text{sign}[x - x_1(\beta_1, \beta_2)][f_C(x, \beta_1, \beta_2) - g_C(x)] < 0$, where $\text{sign}(\cdot)$ is a sign function of \cdot .
- (iii) For any $x \in (0, 1)$, the cooperator's payoff increases in β_2 and the defector's payoff increases in β_1 but decreases in β_2 , i.e., $f_{C\beta_2}(x, \beta_1, \beta_2) > 0$, $f_{D\beta_1}(x, \beta_1, \beta_2) > 0$, and $f_{D\beta_2}(x, \beta_1, \beta_2) < 0$; where $f_{i\beta_j} = \partial f_i / \partial \beta_j, i = C, D, j = 1, 2$. However, the cooperator's payoff increases in β_1 when the group cooperation rate x is higher than $x_2(\beta_1, \beta_2)$ but decreases when x is lower than $x_2(\beta_1, \beta_2)$, i.e., $\text{sign}[x - x_2(\beta_1, \beta_2)][f_{C\beta_1}(x, \beta_1, \beta_2)] > 0$.
- (iv) If $\beta_1 + \beta_2 > 1$, then $x_1(\beta_1, \beta_2) > x_2(\beta_1, \beta_2)$.

Proof. From (18) and (19), for any $\beta_1, \beta_2 \in [0, 1]$ and $x \in (0, 1)$, $f_D(\beta_1, \beta_2, x) < g_D(x)$ and $f_C(x, \beta_1, \beta_2) - g_C(x) = \{- (1 - \beta_1^2)R - [1 - \beta_1(1 - \beta_2)](-S)\}x + [1 - \beta_1(1 - \beta_2)](-S)$. $f_C(x, \beta_1, \beta_2) < g_C(x)$ when $x > x_1(\beta_1, \beta_2)$ and

$f_C(x, \beta_1, \beta_2) > g_C(x)$ when $x < x_1(\beta_1, \beta_2)$. Conclusions (I) and (II) are proved.

Furthermore, one can obtain $f_{C\beta_2}(x, \beta_1, \beta_2) = (1-x)(-S)\beta_1$, $f_{D\beta_1}(x, \beta_1, \beta_2) = (1-\beta_2)Tx$, and $f_{D\beta_2}(x, \beta_1, \beta_2) = -\beta_1Tx$. From $f_{C\beta_1}(x, \beta_1, \beta_2) = [2R\beta_1 + (-S)(1-\beta_2)]x + S(1-\beta_2)$, $f_{C\beta_1}(x, \beta_1, \beta_2) > 0$ when $x > x_2(\beta_1, \beta_2)$ and $f_{C\beta_1}(x, \beta_1, \beta_2) < 0$ when $x < x_2(\beta_1, \beta_2)$. Based on $T > R > 0 > S$, if $\beta_1 + \beta_2 > 1$, then $x_1(\beta_1, \beta_2) > x_2(\beta_1, \beta_2)$. Conclusions (III) and (IV) are proved.

By applying Proposition 4, for any $\beta_1, \beta_2 \in (0, 1)$, the TBC increases or decreases the cooperator's payoff depending on the group cooperation rate and $x_1(\beta_1, \beta_2)$. Fortunately, the critical point $x_1(\beta_1, \beta_2)$ satisfies $\partial x_1(\beta_1, \beta_2)/\partial \beta_i > 0 (i = 1, 2)$. Furthermore, we obtain

$$\lim_{\beta_i \rightarrow 1} x_1(\beta_1, \beta_2) = 1. \quad (22)$$

So, the TBC can always improve the cooperator's payoff if cooperation identified rate β_1 is large enough. Furthermore, for any $x \in (0, 1)$, we can design an appropriate β_1 , such that x_1 . Applying (I) and (II) of Proposition 4, an appropriate TBC can raise the cooperator's payoff and reduce the defector's payoff, thus increasing the group cooperation rate rapidly.

According to (III) of Proposition 4, for any $x \in (0, 1)$, with β_2 increasing, the cooperator's payoff increases and defector's payoff reduces. An interesting phenomenon is $f_{D\beta_1}(x, \beta_1, \beta_2) > 0$ for any $x \in (0, 1)$, i.e., a defector's payoff will always increase in β_1 . Proposition 4 shows that the effect of cooperation identified rate β_1 on an individual's payoff is more complex than defection identified rate β_2 . The cooperator's payoff increases in β_1 when the group cooperation rate C is higher than the critical value $x_2(\beta_1, \beta_2)$ but decreases in β_1 when $x < x_2(\beta_1, \beta_2)$. Based on (II), TBC raises the cooperator's payoff only when $x < x_1(\beta_1, \beta_2)$. So, only when $x \in (x_2(\beta_1, \beta_2), x_1(\beta_1, \beta_2))$, TBC promotes the cooperator's payoff, and larger β_1 results in larger cooperator's payoff.

According to (VI), if $\beta_1 + \beta_2 > 1$, $x_2(\beta_1, \beta_2) < x_1(\beta_1, \beta_2)$. Furthermore, $\lim_{\beta_2 \rightarrow 1} x_2(\beta_1, \beta_2) = 0$, i.e., the cooperator's payoff increases in cooperation identified rate β_1 if defection identified rate β_2 is large enough. At the same time, if $\beta_1 > (-S/2R)((1/x) - 1)$, then the cooperator's payoff increases in β_1 for any $\beta_2 \in (0, 1)$. On the other hand, increasing β_1 can also increase the defector's payoff because the probability of the strategy pair (D, C) is $\beta_1(1-\beta_2)x(1-x)$. Thus, increasing β_1 increases the probability that a defector gains the payoff T !

According to (III) of Proposition 4, for a given defection identified rate β_2 , if the group cooperation rate is very low, for example, $x < ((1-\beta_2)(-S))/(2R + (1-\beta_2)(-S))$, then increasing cooperation identified rate β_1 will not only bring harm to cooperators but also help defectors to get more payoffs. A larger β_1 will doubtlessly cause great damage to the group if the defection identified rate β_2 and the group cooperation rate x are all little.

Denoting the group's expected payoff without TBC as $H(x)$ and with TBC as $H'(x, \beta_1, \beta_2)$, we get the following proposition. \square

Proposition 5. *If $S + T > 0$, then, for any given $x \in (0, 1)$ and any $\beta_1, \beta_2 \in (0, 1)$, the group's expected payoff without TBC is higher than that with TBC, i.e.,*

$$H'(x, \beta_1, \beta_2) < H(x). \quad (23)$$

Proof. Applying the payoff matrix (13), we get $H'(x, \beta_1, \beta_2) = x f_C(x, \beta_1, \beta_2) + (1-x) f_D(x, \beta_1, \beta_2)$ and $H(x) = x g_C(x) + (1-x) g_D(x)$. If $S + T > 0$, it yields

$$H'(x, \beta_1, \beta_2) - H(x) = -(1-\beta_1^2)R x^2 - [1-\beta_1(1-\beta_2)](S+T)(x-x^2) < 0. \quad (24)$$

This completes the proof.

Proposition 5 is simple but interesting. It shows that the TBC reduces the group's average payoff when the cooperator's loss $-S$, caused by defection, is less than the defector's gain T . Furthermore, this conclusion is not affected by x , R , β_1 and β_2 . Furthermore, we obtain

$$\frac{\partial [H'(x, \beta_1, \beta_2) - H(x)]}{\partial \beta_1} = [2\beta_1 R - (1-\beta_2)(S+T)]x^2 - (1-\beta_2)(S+T)x. \quad (25)$$

When $\beta_1 + \beta_2 > 1$, based on (3), $2\beta_1 R - (1-\beta_2)(S+T) > 0$. By letting $x_3(\beta_1, \beta_2) \triangleq (((1-\beta_2)(S+T))/(2\beta_1 R + (1-\beta_2)(S+T)))$, we can obtain $\text{sign}\{[x - x_3(\beta_1, \beta_2)](\partial H'(x, \beta_1, \beta_2) - H(x))/\partial \beta_1\} > 0$. This is consistent with the result of Proposition 2. By the fact that $\lim_{\beta_2 \rightarrow 1} x_3(\beta_1, \beta_2) = 0$, when β_2 is large enough, increasing β_1 can improve the group's average payoff, but it is still lower than that without TBC.

We assume that $S + T < 0$. Denoting $x_4(\beta_1, \beta_2) \triangleq \sqrt{(-[1-\beta_1(1-\beta_2)](S+T))/((1-\beta_1^2)R - [1-\beta_1(1-\beta_2)](S+T))}$, we get $\text{sign}[x - x_4(\beta_1, \beta_2)](H'(x, \beta_1, \beta_2) - H(x)) < 0$. In other words, in the situation where two players' total payoff with strategy pair (C, D) is lower than the total payoff with (D, D) , the TBC can increase the group's average payoff for a lower cooperation rate. \square

Remark 1. Propositions 2–5 are all based on the payoff matrix (13) that is obtained by $m \gg 1$ and $P = 0$. The assumption $m \gg 1$ is realistic. For example, confirmation by the supervisory department of, amongst other things, an enterprise's creditor product quality, cannot be eliminated in short term. The values of $x_i(\beta_1, \beta_2)$ will be more complicated if $P \neq 0, j = 1, 2, 3, 4$. But, the hypothesis that $P = 0$ does not result in significant qualitative or directional changes to conclusions of Propositions 2–5. Furthermore, the ESS and EE of the game do not alter by the assumptions $m \gg 1$ and $P = 0$.

Remark 2. Propositions 2–5 are only static analysis of the impact of TBC on an individual's payoff.

Propositions 2–5 show that TBC will bring down a group's payoff unless the cooperation identified rate β_1 and the defection identified rate β_2 are all large enough. From the point of static analysis, it is hard to understand that TBC can promote cooperation. However, from the evolutionary point

of view, without TBC, cooperation will fail in evolution and all benefits based on cooperation will disappear. Now, we discuss the ESS and EE of the Prisoner's Dilemma with TBC.

3.2. Evolution of Cooperation for the Replicator Dynamics

Proposition 6. *Defection is always ESS of the Prisoner's Dilemma for any $\beta_1, \beta_2 \in (0, 1)$.*

Proof. For any state $(x, y) \in \Delta$ and any $\beta_1, \beta_2 \in (0, 1)$, by payoff matrix (13), one can get $u[(x, y), (0, 1)] = \beta_1(1 - \beta_2)xS < 0$ and $u[(0, 1), (0, 1)] = 0$. So, defection is always ESS.

According to Friedman [43], the Prisoner's Dilemma was discussed as being a linear game; the point $(x, y) = (0, 1)$ is an EE of replicator dynamics (14). Based on Proposition 6, a small group of cooperators cannot invade a group of defectors, no matter how high identified rates β_1 and β_2 are. \square

Proposition 7. *Cooperation is an ESS of the Prisoner's Dilemma with TBC, if and only if the following condition is satisfied:*

$$\frac{\beta_1}{1 - \beta_2} > \frac{T}{R}. \quad (26)$$

By [42], Proposition 7 can be obtained.

Based on Propositions 6 and 7, if and only if (26) is satisfied, the Prisoner's Dilemma with TBC is a symmetric coordination game. The pure strategies "cooperation" and "defection" are all ESS. The fixed points $(1, 0)$ and $(0, 1)$ are all EE. Furthermore, the basin of attraction of the point $(1, 0)$ is $(x_5(\beta_1, \beta_2), 1)$, where $x_5(\beta_1, \beta_2) \triangleq (((1 - \beta_2)(-S))/(\beta_1 R - (1 - \beta_2)(T + S)))$. As $\partial x_5(\beta_1, \beta_2)/\partial \beta_i < 0$, $i = 1, 2$, the basin of attraction of cooperation can rise if we increase the cooperation identified rate β_1 or the defection identified rate β_2 . For any $\beta_1 \in (0, 1)$, $\lim_{\beta_2 \rightarrow 1} x_5(\beta_1, \beta_2) = 0$. This shows that a large-enough defection identified rate β_2 can successfully promote the evolution of cooperation for any β_1 . Furthermore, the basin of attraction of cooperation can be sufficiently large through the designing of an appropriate defection identified rate.

If the defection identified rate satisfies $\beta_2 > (T - R)/T$, cooperation can evolve successfully. By denoting $\lim_{\beta_1 \rightarrow 1} x_5(\beta_1, \beta_2) = ((1 - \beta_2)(-S)/(R - (1 - \beta_2)(T + S))) \triangleq x_6(\beta_2)$, the largest basin of attraction of cooperation is $(x_6(\beta_2), 1)$ for given β_2 . Otherwise, if $\beta_2 < (T - R)/T$, defection is the only ESS. If we consider $(T - R)/T$ as the cost-to-benefit ratio of the cooperation, this conclusion is in good agreement with the work of Nowak [44]. In [44], it is shown that direct reciprocity can lead to the evolution of cooperation only if the probability of another encounter between the same two individuals exceeds the cost-to-benefit ratio of cooperation, i.e., $\beta^d > (T - R)/T$, and indirect reciprocity can only promote cooperation if the probability of knowing

someone's reputation exceeds the cost-to-benefit ratio of the cooperation, i.e., $\beta'' > (T - R)/T$.

By Proposition 7, when condition (26) is satisfied, cooperation is the ESS of the Prisoner's Dilemma and the average payoff of an individual in a group full of cooperators is $\beta_1^2 R$. But, without TBC, defection is the only ESS and the average payoff of an individual will be 0! To sum up the abovementioned discussion, from a dynamic point of view, TBC can promote a cooperator's payoff. Further observation shows that defection identified rate β_2 plays a more important role in promoting the evolution of cooperation than cooperation identified rate β_1 , but β_1 determines an individual's payoff when the cooperation rate is high.

Now, we only discuss the cooperator's payoff and evolution of cooperation.

Proposition 8. *Designing TBC with cooperation identified rate β_1 and the defection identified rate β_2 , we obtain the following conclusions:*

- (i) *When condition (26) is satisfied, but $\beta_1 + \beta_2 - \beta_1\beta_2 + (1/2)\beta_1\beta_2^2 < 1$, this TBC can promote the evolution of cooperation, but for given cooperation rate x in the basin of attraction of cooperation, the cooperator's payoff with TBC is lower than its payoff without TBC*
- (ii) *When $R/T > (1 - \beta_2)[1 - \beta_1(1 - \beta_2)]/(\beta_1 + \beta_2 - 1)$, this TBC can not only expel defectors but also increase the cooperator's payoff for $x \in (x_5(\beta_1, \beta_2), x_1(\beta_1, \beta_2))$*

Proof. If condition (26) is satisfied, based on Proposition 7, the TBC can promote the evolution of cooperation. From the condition $\beta_1 + \beta_2 - \beta_1\beta_2 + (1/2)\beta_1\beta_2^2 < 1$, we obtain that $(1 - \beta_2)[1 - \beta_1(1 - \beta_2)]/(\beta_1 + \beta_2 - 1) > 1$ and $x_5(\beta_1, \beta_2) > x_1(\beta_1, \beta_2)$. The basin of attraction of cooperation is $(x_5(\beta_1, \beta_2), 1)$, and the cooperator's payoff with TBC is higher than its payoff without TBC only when the group cooperation rate $f_C(x, \beta_1, \beta_2) - g_C(x)$ is lower than $x_1(\beta_1, \beta_2)$. So, the TBC can expel defectors, but the cooperator's payoff with TBC is lower than its payoff without TBC when $x \in (x_5(\beta_1, \beta_2), 1)$.

If the condition $R/T > (1 - \beta_2)[1 - \beta_1(1 - \beta_2)]/(\beta_1 + \beta_2 - 1)$ is satisfied, we get $\beta_1 + \beta_2 - \beta_1\beta_2 + (1/2)\beta_1\beta_2^2 > 1$ and $x_5(\beta_1, \beta_2) < x_1(\beta_1, \beta_2)$. So, the TBC can not only expel defectors but also increase the cooperator's payoff for $x \in (x_5(\beta_1, \beta_2), x_1(\beta_1, \beta_2))$.

The condition of (II) of Proposition 8 is stricter than (26). So, based on (II) of Proposition 8, the TBC can not only expel defectors if $x \in (x_1(\beta_1, \beta_2), 1)$ but also increase the cooperator's payoff for $x \in (0, x_5(\beta_1, \beta_2))$. But, only when $x \in (x_5(\beta_1, \beta_2), x_1(\beta_1, \beta_2))$, the TBC can result in these two advantages.

Cooperation identified rate β_1 and defection identified rate β_2 are the input of TBC. They are feedback of cooperation rate x . At any time t , cooperation rate is variable $x(t)$. So, β_1 and β_2 are functions of time t . We denote $\beta_1(t)$ and $\beta_2(t)$ as the cooperation identified rate and the

defection identified rate at time t in the next part of this paper. \square

4. The Optimal Identified Rate of Promoting the Cooperation Rate

4.1. *The Optimal Identified Rate Model.* We define $u_i(t)$ as

$$u_i(t) = \ln\left(\frac{1}{\beta_i(t)} - 1\right), \quad i = 1, 2. \quad (27)$$

Based on $\beta_i \in (0, 1)$, one can obtain $u_i \in (-\infty, +\infty)$ and $\beta_i(t) = (1/(1 + e^{u_i(t)}))$, $i = 1, 2$. The performance index is introduced:

$$J = 1 - x(t_f) + \int_0^{t_f} q_1\beta_1(t) + q_2\beta_2(t)dt, \quad (28)$$

where t_f is the terminal time of the replicator dynamic (14), $x(t_f)$ denotes the terminal cooperation rate, $1 - x(t_f)$ is the effective cost of promoting the cooperation rate, and $\int_0^{t_f} \beta_i(t)dt$ is the control cost of the entire TBC process. The bigger $x(t_f)$ corresponds to the better effect of promoting cooperation and the lower effective cost. The parameter $q_i > 0$ is used to adjust the difference between the effective cost and the control cost in the dimension and the weight. Furthermore, $q_i > 0$ reflects the controller's balance of the TBC for cooperation and for defection.

We rewrite replicator dynamic (14) and performance index (28) under transformation (27) as the follows:

$$\dot{x}(t) = \left\{ \left[\frac{1}{1 + e^{u_1}} R - \frac{e^{u_2}}{1 + e^{u_2}} (S + T) \right] x + \frac{e^{u_2}}{1 + e^{u_2}} S \right\} \frac{1}{1 + e^{u_1}} (1 - x)x, \quad (29)$$

$$J = 1 - x(t_f) + \int_0^{t_f} \frac{q_1}{1 + e^{u_1}} + \frac{q_2}{1 + e^{u_2}} dt. \quad (30)$$

We suppose the initial time of replicator dynamic (14) is $t = 0$, and the initial value of system is

$$x(0) = x_0, \quad 0 < x_0 < 1. \quad (31)$$

The initial value (31) indicates that, at the beginning of the evolution, the population is in a state in which cooperation and defection are mixed. We denote $u(t) = (u_1(t), u_2(t))^T$. We consider $u(t)$ as a control variable of the dynamic (29).

The feedback control law of the state $x(t)$ is designed so that the total cost defined by the performance index (30) is minimized. The optimal identified rate problem can be expressed as follows: under dynamic constraint (29) and initial value constraint (31), the optimal control $u^*(t)$ and the optimal state trajectory $x^*(t)$ are designed to minimize (30).

4.2. *The Optimal Identified Rate Designing Process.* There is no analytic solution for the optimal control problem (29)–(31). Now, we give a numerical algorithm to design an optimal identified rate.

Theorem 1. *Consider the optimal identified rate design described by replicator dynamic (14) with performance index (30). Under (27) and (31), the optimal identified rate $\beta_i^*(t)$ is existent and unique, and $\beta_i^*(t)$ is expressed in the following form:*

$$\beta_1^*(t) = \lim_{h \rightarrow \infty} \lim_{k \rightarrow \infty} \frac{q_2}{\lambda^{(h,k)}(t) (1 - x^{(h,k)}) x^{(h,k)} h(x^{(h,k)})},$$

$$\beta_2^*(t) = \lim_{h \rightarrow \infty} \lim_{k \rightarrow \infty} \left[1 + \frac{q_1 h(x^{(h,k)}) - 2Rq_2 x^{(h,k)}}{\lambda^{(h,k)} (1 - x^{(h,k)}) x^{(h,k)} h^2(x^{(h,k)})} \right], \quad (32)$$

where $\lambda^{(h,k)}(t)$ is the solution of the following adjoint variable differential equation:

$$\lambda^{(h,0)}(t) \equiv 1,$$

$$\dot{\lambda}^{(h,k)}(t) = \frac{q_2^2 R [x^{(*)h}]^2 [3h(x^{(*)h}) - 2T] + q_1 q_2 h(x^{(*)h}) [(-3x^{(*)h} + 1)h(x^{(*)h}) + T x^{(*)h}]}{\lambda^{(h,k-1)} [x^{(*)h}]^2 (1 - x^{(*)h})^2 h^3(x^{(*)h})}, \quad (33)$$

$$\lambda^{(h,k)}(t_f) = 1, \quad h, k = 1, 2, \dots,$$

where $x^{(*)0}(t) = x_0$, $x^{(*)h}(t) = \lim_{k \rightarrow \infty} x^{(h,k)}(t)$, and $x^{(h,k)}(t)$ is the solution of the following state differential equation:

$$x^{(h,0)}(t) \equiv x_0,$$

$$\lambda^{(*_0)}(t) \equiv 1,$$

$$\dot{x}^{(h,k)}(t) = \frac{q_1 q_2 h(x^{(h,k-1)}) - R q_2^2 x^{(h,k-1)}}{\left[\lambda^{(*_{h-1})} \right]^2 x^{(h,k-1)} (1 - x^{(h,k-1)}) h^2(x^{(h,k-1)})},$$

$$t \in [0, t_f],$$

$$x^{(h,k)}(0) = x_0, \quad h, k = 1, 2, \dots,$$

(34)

where $\lambda^{(*_h)}(t) = \lim_{k \rightarrow \infty} \lambda^{(h,k)}(t)$.

Proof. By introducing the adjoint variable $\lambda(t)$ and letting $h(x) = (S + T)x - S$, if the optimal control and the optimal state are $u^*(t)$ and $x^*(t)$, then there exists the adjoint variable $\lambda^*(t)$ such that $x^*(0) = x_0$, $\lambda^*(t_f) = 1$, and

$$q_1 + \lambda(t)(1-x)x \left[-\frac{2Rx}{1+e^{u_1}} + \frac{e^{u_2}}{1+e^{u_2}} h(x) \right] \Big|_{\lambda=\lambda^*, x=x^*, u=u^*} = 0, \quad (35)$$

$$q_2 - \lambda(t)(1-x)xh(x) \frac{1}{1+e^{u_1}} \Big|_{\lambda=\lambda^*, x=x^*, u=u^*} = 0. \quad (36)$$

From (36), we obtain $x^*(t)$, $\lambda^*(t)$, and $u_i^*(t)$ satisfying the following equations:

$$\frac{1}{1+e^{u_1}} \Big|_{\lambda=\lambda^*, x=x^*, u=u^*} = \frac{q_2}{\lambda(t)(1-x)xh(x)} \Big|_{\lambda=\lambda^*, x=x^*, u=u^*}, \quad (37)$$

$$\frac{1}{1+e^{u_2}} \Big|_{\lambda=\lambda^*, x=x^*, u=u^*} = \left(1 + \frac{q_1}{\lambda(t)(1-x)xh(x)} - \frac{2Rq_2x}{\lambda(t)(1-x)xh^2(x)} \right) \Big|_{\lambda=\lambda^*, x=x^*, u=u^*}. \quad (38)$$

The adjoint variable and its terminal condition with the optimal control $u^*(t)$ are as follows:

$$-\dot{\lambda}(t) = \lambda(t) \left\{ \frac{Rx(2-3x)}{(1+e^{u_1})^2} - \frac{e^{u_2}}{1+e^{u_2}} \frac{1}{1+e^{u_1}} [(S \right.$$

$$+ T)x(2-3x) - S(1-2x)] \},$$

$$\lambda(t_f) = 1. \quad (40)$$

Substituting (37) and (38) into (29) and (39) yields the following two-point boundary value (TPBV) problem:

$$\dot{x}(t) = \frac{q_1 q_2 h(x) - R q_2^2 x}{\lambda^2 x (1-x) h^2(x)},$$

$$-\dot{\lambda}(t) = \frac{q_2^2 R x^2 [3h(x) - 2T] + q_1 q_2 h(x) [(-3x + 1)h(x) + Tx]}{\lambda x^2 (1-x)^2 h^3(x)},$$

$$x(0) = x_0,$$

$$\lambda(t_f) = 1.$$

(41)

Note that the TPBV problem (41) is nonlinear; for a nonlinear TPBV problem, there is no analytic solution. From (37), (38), and (27), one can obtain

$$\beta_1^*(t) = \frac{q_2}{\lambda^*(t)(1-x^*)x^*h(x^*)},$$

$$\beta_2^*(t) = 1 + \frac{q_1}{\lambda^*(1-x^*)x^*h(x^*)} - \frac{2Rq_2x^*}{\lambda^*(1-x^*)x^*h^2(x^*)}, \quad (42)$$

where $\lambda^*(t)$ and $x^*(t)$ is the solution to the TPBV problem (41). A sequence of state differential equations is constructed:

$$x^{(1,0)}(t) \equiv x_0,$$

$$\dot{x}^{(1,k)}(t) = \frac{q_1 q_2 h(x^{(1,k-1)}) - R q_2^2 x^{(1,k-1)}}{x^{(1,k-1)} (1 - x^{(1,k-1)}) h^2(x^{(1,k-1)})}, \quad t \in [0, t_f],$$

$$x^{(1,k)}(0) = x_0, \quad k = 1, 2, \dots$$

(43)

For certain k , $x^{(1,k-1)}(t)$ is known, so equation (43) is a sequence of differential equations. Furthermore, $x^{(1,k)}(t)$ can be solved from (43). According to Tang and Dong [45], the solution sequence defined by (43) is uniformly convergent to the solution of the following equation:

$$\dot{x}(t) = \frac{q_1 q_2 h(x) - R q_2^2 x}{x(1-x)h^2(x)}, \quad t \in [0, t_f], \quad (44)$$

$$x(0) = x_0.$$

We define $\lambda^{(*_1)}(t) = \lim_{k \rightarrow \infty} x^{(1,k)}(t)$ and construct the following sequence:

$$\lambda^{(1,0)}(t) \equiv 1,$$

$$\dot{\lambda}^{(1,k)}(t) = \frac{q_2^2 R \left[x^{(*)} \right]^2 \left[3h(x^{(*)}) - 2T \right] + q_1 q_2 h(x^{(*)}) \left[(-3x^{(*)} + 1)h(x^{(*)}) + Tx^{(*)} \right]}{\lambda^{(1,k-1)} \left[x^{(*)} \right]^2 \left(1 - x^{(*)} \right)^2 h^3(x^{(*)})}, \quad (45)$$

$$\lambda^{(1,0)}(t_f) = 1, \quad h, k = 1, 2, \dots$$

In the k th iteration, $\lambda^{(1,k-1)}(t)$ and $f_{C_x}(x, \beta_1, \beta_2) - g_{C_x}(x) < 0$ are known, so (45) is also a sequence of differential

equations. Based on [45], sequence (45) is uniformly convergent to the solution of the following equation:

$$\dot{\lambda}(t) = \frac{q_2^2 R \left[x^{(*)} \right]^2 \left[3h(x^{(*)}) - 2T \right] + q_1 q_2 h(x^{(*)}) \left[(-3x^{(*)} + 1)h(x^{(*)}) + Tx^{(*)} \right]}{\lambda \left[x^{(*)} \right]^2 \left(1 - x^{(*)} \right)^2 h^3(x^{(*)})}, \quad (46)$$

$$\lambda(t_f) = 1.$$

We denote $\lambda^{(*)}(t) = \lim_{k \rightarrow \infty} \lambda^{(1,k)}(t)$. By constructing sequences (33) and (34) and repeating the previous iteration process, we can obtain the sequences $\{x^{(h,k)}(t)\}$ and $\{\lambda^{(h,k)}(t)\}$. The sequences $\{x^{(h,k)}(t)\}$ and $\{\lambda^{(h,k)}(t)\}$ are uniformly convergent to the solution of TPBV (39), i.e.,

$$\begin{aligned} \lim_{h \rightarrow \infty} \lim_{k \rightarrow \infty} \lambda^{(h,k)}(t) &= \lambda^*(t), \\ \lim_{h \rightarrow \infty} \lim_{k \rightarrow \infty} x^{(h,k)}(t) &= x^*(t). \end{aligned} \quad (47)$$

This completes the proof.

The number of iterations X and Y can be determined by the accuracy requirement for the performance index (30). We give a practical algorithm calculating the (X, Y) th suboptimal optimal identified rate as Algorithm 1 \square

5. Simulation Examples

In the Prisoner's Dilemma with payoff matrix (1), let $t_f = 5$, $T = 5$, $R = 3$, $P = 0$, $S = -4$. The replicator dynamic with TBC is

$$\begin{aligned} \dot{x}(t) &= \{(3\beta_1(t) + \beta_2(t) - 1)x(t) - 4(1 - \beta_2(t))\}\beta_1(t) \\ &\quad - x(t)x(t), \quad t \in [0, 5]. \end{aligned} \quad (48)$$

The replicator dynamic without TBC is

$$\dot{x}'(t) = 2(x'(t) - 2)(1 - x'(t))x'(t), \quad t \in [0, 5]. \quad (49)$$

By simulation, we study the following three problems.

5.1. Effect Analysis of TBC. Let $x_0 = 0.4$, $\beta = \beta_1(t) = \beta_2(t) = 0.7$ and 0.9 , and $t \in [0, 5]$, respectively. In the previous discussion, we denote $f_C(x, \beta_1, \beta_2)$ as the payoff of the cooperator with TBC, denote $f_D(x, \beta_1, \beta_2)$ as the payoff of

the defector with TBC, denote $g_C(x)$ as the payoff of the cooperator without TBC, and denote $g_D(x)$ as the payoff of the defector without TBC, denoting the group's expected payoff without TBC as $H(x)$ and with TBC as $H'(x, \beta_1, \beta_2)$. In Figure 1, to simplify legends and discuss the effect of β on promoting cooperation, we redefine the symbol as follows: at $t = t_0$, let $H'(\beta)$ and $f_i(\beta)$ be the payoffs of group and individual i with TBC β from $t = 0$ to $t = t_0$, respectively, $H(\beta)$ and $g_i(\beta)$ be obtained after TBC has just been withdrawn, i.e., TBC β is used from $t = 0$ to $t = t_0 - \tau$, where τ is positive and small enough, $x(\beta)$ is the cooperation rate with TBC β , g_i is the payoff function curve of individual i , H is the payoff function curve of the group, and x is the cooperation rate curve without TBC, respectively, where $i = C, D$.

Based on Figures 1(a) and 1(b), from the curves of g_C and g_D , we can know the payoffs of the cooperator and defector are all decreasing and cooperation fails to evolve without TBC. Comparing Figures 1(a) and 1(d), we can see that cooperators obtain more payoff than without TBC, but cooperation rate $x(0.7)$ decreases and cooperators are expelled still because the identified rate is low when $\beta = 0.7$. According to the parameters in this example, $x_0 = 0.4$ is in the basin of attraction of cooperation only when $\beta > 11/14$. So, as a result, payoff of cooperator decreases and cooperation fails to evolve. When $\beta = 0.9$, payoff of cooperators increases and cooperation evolves successfully; payoff of defectors increases but defection fails to evolve.

Here, we try to show what will happen if we withdraw TBC for a short time. First of all, we emphasize the symbols $H'(0.9)$, $H(0.9)$, and H of Figure 1(c). For a given time t_0 , $H(0.9)$ and $H'(0.9)$ all correspond to the cooperation rate $x(t_0)$ defined by the replicator dynamics (48) with TBC $\beta = 0.9$, but H corresponds to the cooperation rate $x'(t_0)$ defined by the replicator dynamics without TBC (49). As shown in Figure 1(b), $g_D(\beta) < f_D(\beta)$ for any identified rate

Input: Given positive real numbers δ and ε .

Output: (X, Y) th suboptimal $\beta_i^{(*,x)}(t)$, $i = 1, 2$.

Step 1: Initialize $h = 1$, $\lambda^{(*,0)}(t) = 1$.

Step 2: Initialize $k = 1$, $x^{(h,0)}(t) = x_0$.

Step 3: Calculate (h, k) th state variable $x^{(h,k)}(t)$ according to equation (34).

Step 4: If $\|x^{(h,k)}(t) - x^{(h,k-1)}(t)\| < \delta$, let $x^{(*,h)}(t) = x^{(h,k)}(t)$; otherwise, let $k = k + 1$ and jump to Step 3.

Step 5: Initialize $k = 1$, $\lambda^{(h,0)}(t) = 1$.

Step 6: Calculate (h, k) th adjoint variable $\lambda^{(h,k)}(t)$ according to equation (33).

Step 7: If $\|\lambda^{(h,k)}(t) - \lambda^{(h,k-1)}(t)\| < \delta$, let $\lambda^{(*,h)}(t) = \lambda^{(h,k)}(t)$; Otherwise, let $k = k + 1$ and jump to Step 6.

Step 8: Letting $X = h$, calculate .. according to the following equation:

$$J^{(H)} = 1 - x(t_f) + \int_0^{t_f} (q_1 \beta_1^{(*,H)} + q_2 \beta_2^{(*,H)}) dt$$

Step 9: If $|(J^{(X)} - J^{(X-1)})/J^{(X)}| < \varepsilon$ stop the algorithm and calculate the suboptimal $\beta_i(t)$, $i = 1, 2$ according to equation (32); otherwise, let $h = h + 1$ and jump to Step 2.

ALGORITHM 1: Finding a suboptimal optimal identified rate.

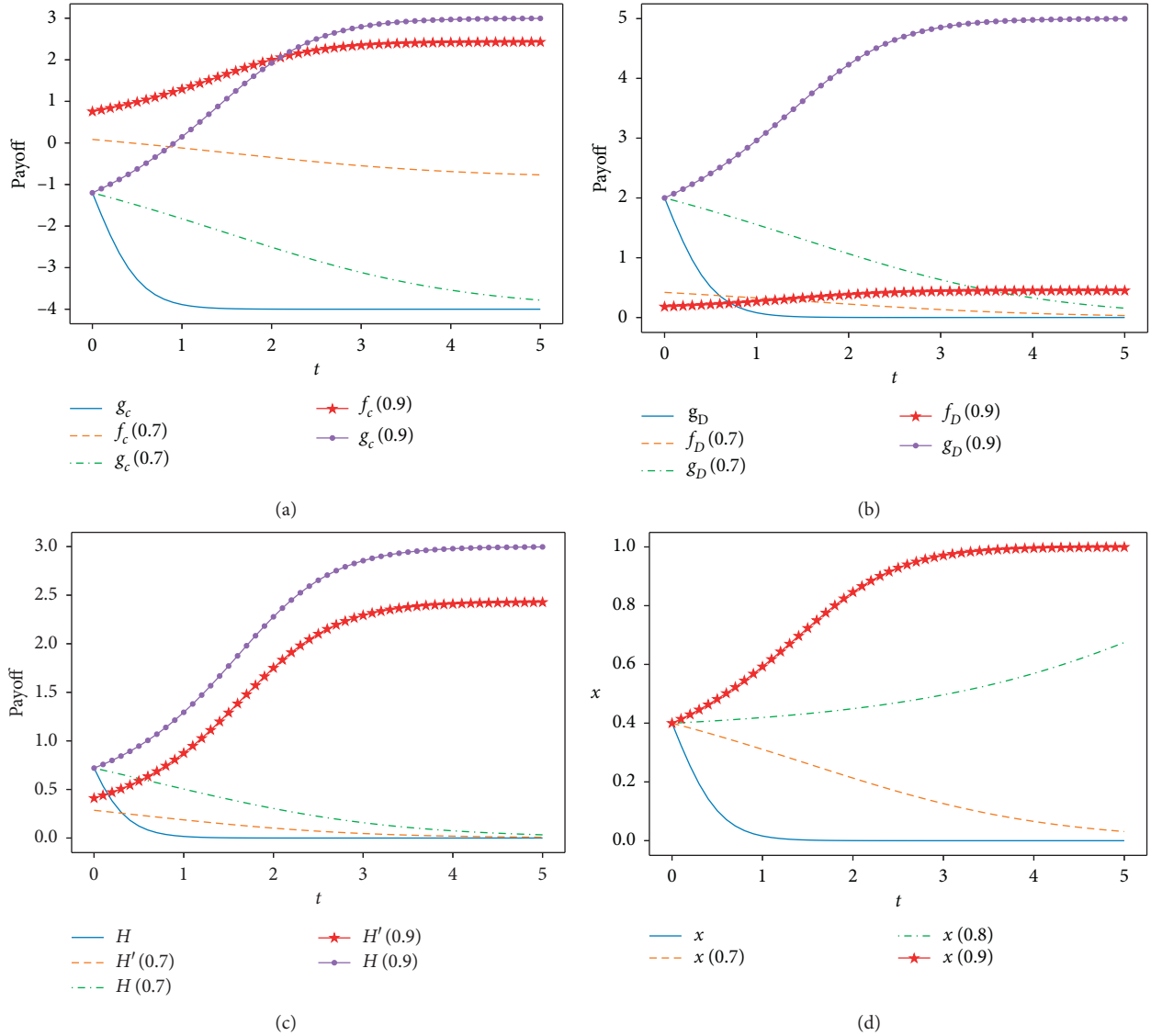


FIGURE 1: (a). The payoffs of the cooperator. (b). The payoffs of the defector. (c). The payoff of the group. (d). Evolution of the cooperation under TBC.

β . When $\beta = 0.7$, condition (26) is not satisfied and $x = 0.861$, so the point $x_0 = 0.4$ is not in the basin of attraction of cooperation and $f_C(0.7) > g_C(0.7)$. But, by designing a large identified rate $\beta = 0.9$, we obtain the basin of attraction of cooperation is $(0.154, 1)$ and $x_1 = 0.861$. When $t > 2.213$, one gets $x > 0.861$ and $f_C(0.9) > g_C(0.9)$ so that TBC inhibits the growth rate of the cooperator's payoff. If we withdraw TBC for a short time, cooperators will obtain more payoffs.

Combined with the evolution of cooperation, we come to the following conclusion: *when the cooperation rate $x(t)$ is low, TBC results in the payoff for all individuals being higher than the payoff without TBC; when the cooperation rate $x(t)$ is large enough, the payoff for all individuals will be lower than the payoff without TBC under same $x(t)$.* On one hand, Figures 1(b) and 1(c) show that TBC decreases not only the defector's payoff but also the group's payoff. Furthermore, the higher the cooperation rate is, the more obvious the inhibition is. On the other hand, we observe the positive effects of TBC on promoting both the cooperation rate and the group's payoff. Only under a bigger $\beta_i(t)$, $i = 1, 2$ does the cooperation rate $x(t)$ increase. For example, although the group payoff $H'(0.9)$ under TBC is always smaller than $H(0.9)$, it is always greater than H .

In summary, Figure 1 shows that when the group cooperation rate $x(t)$ defined by (48) reaches a satisfactory value, temporary withdrawal of TBC can increase the payoff of all individuals in the group for a short time. However, if without TBC for a long time, the group cooperation rate $\beta_i(t)$, $i = 1, 2$ defined by (49) will drop to a very low level and all individual benefits will decrease.

5.2. Analysis of Rationality of the Optimal Identified Rate. Letting $x(0) = 0.4$, $q_1 = q_2 = q$, we obtain the optimal identified rate $\beta_i^*(t)$, $i = 1, 2$ under different q as Figure 2 and the optimal state path as Figure 3.

Given a sufficiently large constant $\Gamma > 0$ and a sufficiently small constant $\gamma > 0$, if $u_i(t) > \Gamma$, we define $\beta_i(t) = 0$, and if $0 < u_i(t) < \gamma$, we define $\beta_i(t) = 1$, where $i = 1, 2$. Figure 2 shows that $\beta_i^*(t)$, $i = 1, 2$ approximates bang-bang control with switching time t_{is} . If $t < t_{is}$, then $\beta_i^*(t) = 1$; otherwise, $\beta_i^*(t) = 0$. The smaller q corresponds to the bigger switching time. If q_i is large enough, then $\beta_i^*(t) = 0$ and $x^*(t) = x_0$, $t \in [0, t_f]$.

Now, we explain $\beta_i^*(t) = 0$. According to the definition of TBC, $\beta_i^*(t) = 0$ means all cooperators are labeled as "defectors" and all defectors are labeled as "cooperators". This is impossible and inefficient. Here, $\beta_i^*(t) = 0$ can be understood as no effective TBC; all individuals choose "defection" strategy for self-protection. In this case, the control cost is 0. Under $\beta_i(t) = 0$, the cooperation rate $x(t)$ is a constant value. By designing the bang-bang control as Figure 2, we can achieve the goal of successfully promoting the evolution of cooperation. This meets the requirement of minimizing the performance index (30). This is, however, unreasonable for the group in terms of revenue. Probably, the more reasonable approach would

be to withdraw TBC when the cooperation rate $x(t)$ is satisfactory. Now, let us discuss the TBC design from the perspective of revenue.

Figure 3 shows that there is a suitable TBC to make the cooperation evolve successfully. Furthermore, the smaller q results in the better effect of promoting cooperation.

By designing $\beta_i^*(t)$ defined by Theorem 1, the corresponding expressions of the total payoff of the cooperator, defector, and group is $\int_0^5 \{ [3\beta_1^2(t) + 4\beta_1(t)(1 - \beta_2(t))]x(t) - 4\beta_1(t)(1 - \beta_2(t)) \} dt$, $\int_0^5 5\beta_1(t)(1 - \beta_2(t))x(t) dt$, and $\int_0^5 [\beta_1(t)(3\beta_1(t) + \beta_2(t) - 1)x^2(t) + \beta_1(t)(1 - \beta_2(t))x(t)] dt$, respectively. If we do not design the TBC, from $t = 0$ to $t = 5$, the total payoffs of the cooperator, defector, and group can be expressed as $\int_0^5 (7x'(t) - 4) dt$, $\int_0^5 5x'(t) dt$, and $\int_0^5 (2x'^2(t) + x'(t)) dt$, respectively, where $x'(t)$ is described by (49). Let $q_1 = q_2 = 0.1$ and $q_1 = 0.1, q_2 = 0.05$, respectively. The calculation results are shown in Table 1.

Table 1 shows that $\beta_i^*(t)$ improves the cooperator's payoff. Furthermore, as q_i decreases, the switching time t_{is} moves backwards and the cooperator's payoff increases.

5.3. Optimal Timing Control (OTC) of Switched Replicator Dynamics. System (48) and (49) is a switched nonlinear system. In the past decades, switched systems have received significant research attention, see, e.g., [46–48]. In this paper, we design the switching control (SC) as follows: for a given q_1 and q_2 , let $t_s = t_{1s} = t_{2s}$, and we design TBC with the identified rate as

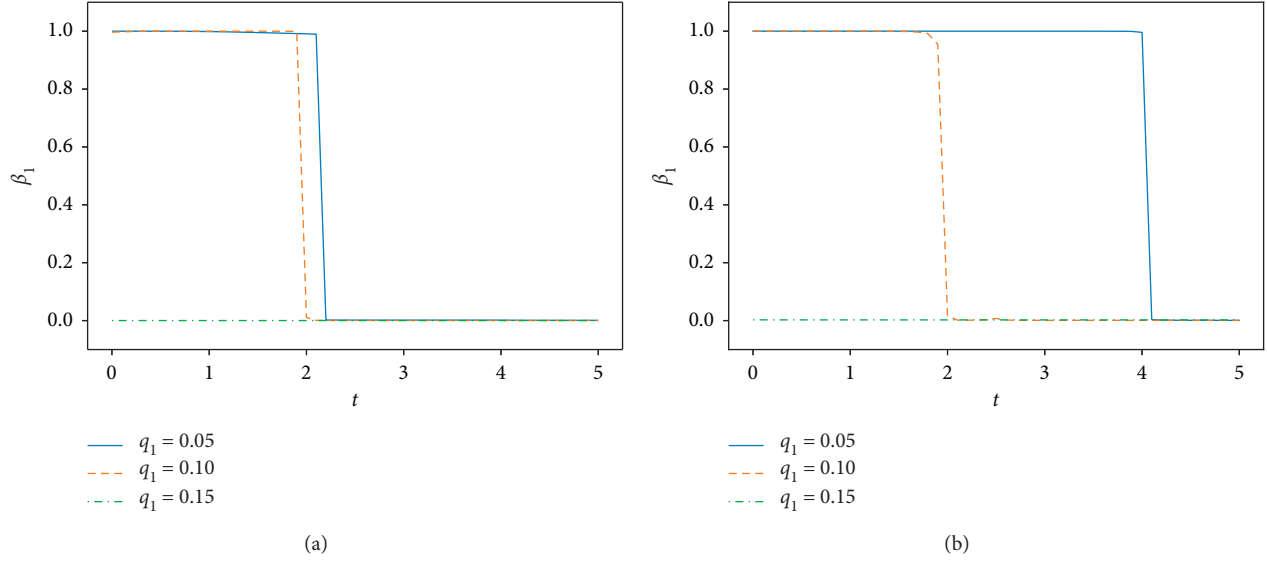
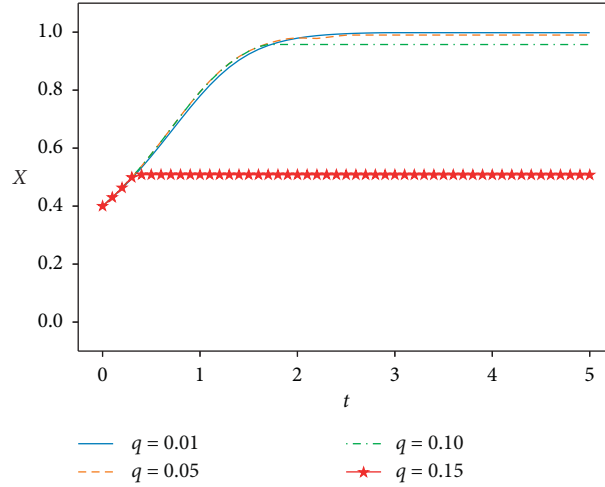
$$\beta_i(t) = 1, \quad t \in [0, t_s], i = 1, 2. \quad (50)$$

However, when $t > t_s$, we give up the TBC. The replicator dynamics with SC is as follows:

$$\begin{aligned} \dot{x}''(t) &= \begin{cases} 3(1 - x''(t))x''^2(t), & t \in [0, t_s], \\ (x''(t) - 2)(1 - x''(t))x''(t), & t \in [t_s, 5], \end{cases} \\ x''(0) &= x_0, \\ \lim_{t \rightarrow t_s^-} x''(t) &= x''(t_s). \end{aligned} \quad (51)$$

Assuming $x_0 = 0.4$, the switching time t_s is decided by the optimal identified rate problem (29)–(31) according to different weight coefficients q_i . Based on a switched system (51), the total payoffs of the cooperator, defector, and group can be expressed as $\int_0^{t_s} 3x''(t) dt + \int_{t_s}^5 (7x''(t) - 4) dt$, $\int_{t_s}^5 5x''(t) dt$, and $\int_0^{t_s} 3x''^2(t) dt + \int_{t_s}^5 (2x''^2(t) + x''(t)) dt$, respectively. The simulation results are shown in Table 2, and the state path with SC is shown in Figure 4.

By comparing Tables 1 and 2, we find the impact of control on the average payoff of a group is complex. Table 1 shows that the smaller coefficient q_i corresponds to a large group's payoff under optimal TBC, but Table 2 shows the opposite result. Combining with Figures 3 and 4, the group

FIGURE 2: The optimal identified rate $\beta^*(t)$ under different q .FIGURE 3: Cooperation rate $x^*(t)$ with $\beta^*(t)$.

has a larger average payoff under SC, but the cooperation rate is lower at $t = t_f = 5$.

So, how to choose an appropriate control, TBC or SC, and how to choose the switching time t_s depend on what we pay more attention to.

Now, we design the cost function J_{SC} with OTC as follows:

$$J_{SC} = x''(5) - 2qt_s + q' \left[\int_0^{t_s} 3x''(t)dt + \int_{t_s}^5 (7x''(t) - 4)dt \right], \quad (52)$$

where $x(5)$ is the terminal cooperation rate, $2t_s$ is the control cost, $\int_0^{t_s} 3x''(t)dt + \int_{t_s}^5 (7x''(t) - 4)dt$ is the cooperator's average payoff, and $q > 0$ is its weight coefficient. We search the optimal switching time t_s to maximize the function (52), i.e., we design an OTC for promoting evolution of the Prisoner's Dilemma and increasing the cooperator's average payoff.

Theorem 2. Consider the OTC described by the switched replicator dynamic (51) with performance index (52). $\varphi_1(t_s)$ and $\varphi_2(t, t_s)$ are solutions to algebra equation (53) and implicit function (54), respectively:

$$\frac{\varphi_1(t_s)}{1 - \varphi_1(t_s)} e^{-(1/\varphi_1(t_s))} = \frac{x_0}{1 - x_0} e^{-(1/x_0)} e^{3t_s}, \quad (53)$$

$$\frac{1}{(2 - \varphi_2(t, t_s)\varphi_2(t, t_s))} = 1 + \frac{(1 - \varphi_1(t_s))^2}{(2 - \varphi_2(t, t_s)\varphi_1(t_s))} e^{2(t-t_s)}, \quad t \in (t_s, 5]. \quad (54)$$

The optimal switching time t_s is decided by

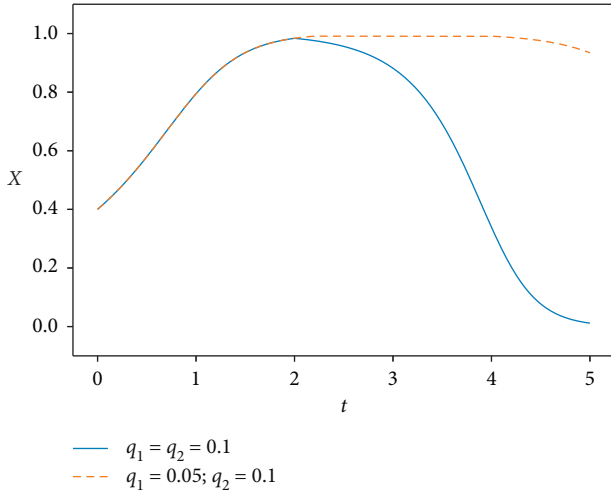
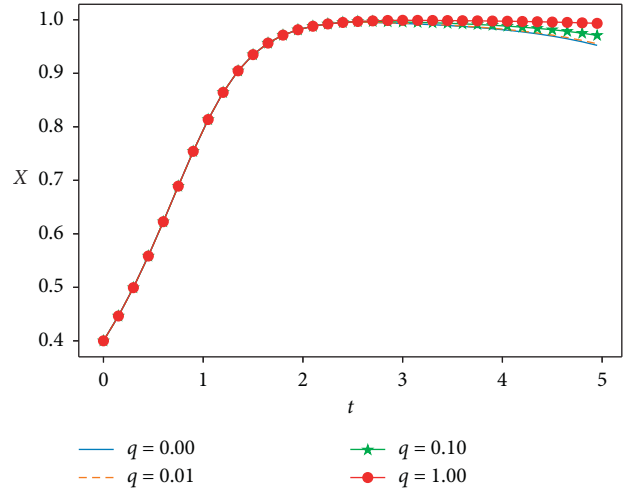
$$\frac{dJ_{SC}}{dt_s} = \frac{d\varphi_2(5, t_s)}{dt_s} - 2q' - 4q\varphi_1(t_s) + 7q \int_{t_s}^5 \frac{\partial \varphi_2(t, t_s)}{\partial t_s} dt = 0. \quad (55)$$

TABLE 1: Payoffs of the cooperator, defector, and group under different q_i

	Cooperator's payoff	Defector's payoff	Group's payoff
$\beta_i^*(t)$ with $q_1 = q_2 = 0.1$	4.581	0	3.618
$\beta_i^*(t)$ with $q_1 = 0.1, q_2 = 0.05$	5.111	0	4.203
No TBC	-18.993	0.719	0.208

TABLE 2: The payoff of the cooperator, defector, and group with SC.

	Cooperator's payoff	Defector's payoff	Group's payoff
SC with $q_1 = 0.1, q_2 = 0.2$	4.571	8.610	8.111
SC with $q_1 = 0.1, q_2 = 0.05$	7.916	4.858	7.067

FIGURE 4: Cooperation rate $x(t)$ with SC.FIGURE 5: Cooperation rate $x(t)$ with OTC.

Proof. The following two implicit functions can be obtained from the switched replicator dynamics (51):

$$\frac{x_2(t)}{1-x_2(t)} e^{-(1/x_2(t))} = \frac{x_0}{1-x_0} e^{-(1/x_0)} e^{3t}, \quad t \in [0, t_s], \quad (56)$$

$$\frac{1}{(2-x_2(t))x_2(t)} = 1 + \frac{(1-x_2(t_s))^2}{(2-x_2(t_s)x_2(t_s))} e^{2(t-t_s)}, \quad t \in (t_s, 5]. \quad (57)$$

Based on existence theorem of an implicit function, we get $x_2(t) = \begin{cases} \varphi_1(t), & t \in [0, t_s) \\ \varphi_2(t, t_s), & t \in [t_s, 5] \end{cases}$ from (56) and (57). So, $\varphi_1(t)$, $t \in [0, t_s)$ satisfies (53) and $\varphi_2(t, t_s)$, $t \in [t_s, 5]$ satisfies (54). The cost function $x(t) + y(t) = 1$ can be rewritten as follows:

$$J_{SC} = \varphi_2(5, t_s) - 2qt_s + q' \left[\int_0^{t_s} 3\varphi_1(t) dt + \int_{t_s}^5 7\varphi_2(t, t_s) dt \right]. \quad (58)$$

Because $d \int_0^{t_s} \varphi_1(t) dt / dt_s = \varphi_1(t_s)$, $d \int_{t_s}^5 \varphi_2(t, t_s) dt / dt_s = -\varphi_2(t_s, t_s) + \int_{t_s}^5 \partial \varphi_2(t, t_s) / \partial t_s dt$, and $\varphi_2(t_s, t_s) = \varphi_1(t_s)$, the optimal switching time t_s is decided by (55).

Letting $q' = 0.1$, $q_1 = 0, 0.01, 0.1$ and 1 , respectively, we get the cooperation rate curve as in Figure 5 and the payoffs of the cooperator, defector, and group as in Table 3 by applying Theorem 2.

Figure 5 shows that the cooperation rate $x(t)$ is very high whenever the weight coefficient of the cooperator's payoff $q = 0, 0.01, 0.1$ or $q = 1$ although the switching time t_s changes (t_s varies from 2.472 to 2.967). It is because we choose a large coefficient of cooperation rate at terminal time $x(5)$. In fact, the main goal of this paper is to promote the evolution of cooperation by designing TBC. So, we design the cost function J_{SC} to promote the cooperation rate and to reconcile control cost and the cooperator's payoff. Observing Table 3, we note that the switching time t_s increases as the cooperator's payoff weight coefficient q gets bigger.

The OTC has the following advantages when Table 3 is compared to Tables 1 and 2. Firstly, it has the same advantage as TBC: it restrains the payoffs of defectors. Secondly, it greatly enhances the payoffs of cooperators compared to TBC. Thirdly, it is the optimal way to promote the payoffs not only for cooperators but also for groups. Lastly, it can effectively promote the cooperation rate $x(t)$. \square

TABLE 3: Switching time and payoffs of individuals under OTC.

	Switching time t_s	Cooperator's payoff	Defector's payoff	Group's payoff
$q = 0$	2.472	13.180	12.409	13.275
$q = 0.01$	2.488	13.201	12.343	13.289
$q = 0.1$	2.597	13.308	12.878	13.364
$q = 1$	2.967	13.458	10.135	13.471

6. Conclusions

The analysis in this paper provides value insights into a number of issues regarding some clusters in the economic society when faced with the Prisoner's Dilemma, which can help operators of clusters decide to adopt TBC or not, identify cooperators or identify defectors, and design the value of the identified rate.

First, operators should judge the harm of defection in the Prisoner's Dilemma. We show that when the cooperation rate is high, to further increase cooperation rate, a large identified rate is needed. Thus, it is difficult to design a TBC to further promote cooperation and further improve the cooperator's payoff. Furthermore, the cluster must pay for the TBC. So, it is not necessary to use TBC for a cluster when the cooperation rate is high. If operator considers TBC is necessary, we advice he/she depends more on the cooperation identification.

By contrast, if the cooperation rate is too low, i.e., the group is almost full of defectors, it is difficult to design a TBC which can help cooperators invade the group of defectors successfully. In this case, any TBC all can increase the cooperator's payoff. But, a large defection identified rate is essential to promote the cooperation rate. It is very important to note that defection identification is more effective than cooperation identification.

We show the switch between adopting TBC and withdrawing TBC is necessary. TBC has an advantage in promoting evolution of cooperation. However, when the cooperation rate is high, TBC reduces the payoffs of the cooperators unless the identified rate is high enough. This paper gives OTC as a way to design switching law. The simulation results confirm OTC not only maintains a high cooperation rate but also increases the payoff of the cooperators.

While this paper is a significant step in the economic analysis of the Prisoner's Dilemma and in the optimal control of promoting cooperation, there are a number of interesting directions for future work in this area. There are a number of potential interesting extensions of our model through relaxation of some of our assumptions. For example, we assume the payoff of mutual defection is 0 in this paper, i.e., $P = 0$. In fact, we can assume an individual is exclusive [49] and every individual only accepts an encounter-tagged "cooperator".

Last but not the least, this paper does not really design an SC law for the switched replicator dynamics; the SC designed in this paper is only an OTC with one switching time. Further discussions for promoting cooperation and increasing the cooperator's average payoff will aim to design an OTC with multiple switching times and general SC

unrestrained by bang-bang control. Furthermore, the optimal tag-based cooperation control proposed in this paper can be applied in management practice, such as in e-commerce platforms. Because of information asymmetry about product quality in e-commerce platforms, there exists the PDG in seller groups and consumer groups. Platform manager can design tag-based cooperation control by offering an authentication service. A valuable extension of this research is to empirically examine the feasibility of TBC proposed in this paper.

Data Availability

The simulation model and data are given in the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This paper was supported by the National Natural Science Foundation of China (71871171, 61773156, 71871173, and 71801175). This work was also partly supported by the Key Scientific Research Project in Universities of Henan Province (19A120006).

References

- [1] V. C. Wynne-Edwards, *Animal Dispersion in Relation to Social Behavior*, Hafner Publishing Company, New York, NY, USA, 1962.
- [2] W. D. Hamilton, "The genetical evolution of social behaviour. II," *Journal of Theoretical Biology*, vol. 7, no. 1, pp. 17–52, 1964.
- [3] B. Allen and M. A. Nowak, "Games among relatives revisited," *Journal of Theoretical Biology*, vol. 378, no. 8, pp. 103–116, 2015.
- [4] J. W. Michael, "Reviewed work: the selfish gene by richard dawkins," *Evolution*, vol. 32, no. 1, pp. 220–221, 1978.
- [5] K. Kobayashi, "Conditions for kin selection to bring cooperation and improve population growth: a response to Iri-tani," *Journal of Ecology*, vol. 108, no. 1, pp. 14–16, 2020.
- [6] J. M. Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. 5427, pp. 15–18, 1973.
- [7] E. Tornero, J. F. Sánchez-Romera, J. J. Morosoli, A. Vázquez, Á. Gómez, and J. R. Ordoñana, "Altruistic behavior among twins," *Human Nature*, vol. 29, no. 1, pp. 1–12, 2018.
- [8] R. L. Trivers, "The evolution of reciprocal altruism," *The Quarterly Review of Biology*, vol. 46, no. 1, pp. 35–57, 1971.

- [9] F. Dercole, F. Della Rossa, and C. Piccardi, "Direct reciprocity and model-predictive rationality explain network reciprocity over social ties," *Scientific Reports*, vol. 9, no. 1, 2019.
- [10] X.-D. Zheng, C. Li, J.-R. Yu et al., "A simple rule of direct reciprocity leads to the stable coexistence of cooperation and defection in the Prisoner's Dilemma game," *Journal of Theoretical Biology*, vol. 420, no. 5, pp. 12–17, 2017.
- [11] R. Axelrod and W. D. Hamilton, "The evolution of cooperation," *Science*, vol. 211, no. 4489, pp. 135–160, 1981.
- [12] M. Milinski, "TIT FOR TAT in sticklebacks and the evolution of cooperation," *Nature*, vol. 325, no. 6103, pp. 433–435, 1987.
- [13] X. Lei, C. Y. Sun, D. Wunsch et al., "An adaptive strategy via reinforcement learning for the prisoner's dilemma game," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 1, pp. 304–313, 2018.
- [14] M. Jia, Y. Xiang, and Z. Zhang, "Indirect reciprocity and corporate philanthropic giving: how visiting officials influence investment in privately owned Chinese firms," *Journal of Management Studies*, vol. 56, no. 2, pp. 372–407, 2019.
- [15] H. Chen, R. Suzuki, and T. Arita, "The evolution of cooperation based on indirect reciprocity and spatial locality in continuous space," *Artificial Life and Robotics*, vol. 3, p. 7649, 2020.
- [16] Y. L. Zhang, A. Z. Liu, and C. Y. Sun, "Development of several studies on indirect reciprocity and the evolution of cooperation," *Acta Automat Sinica*, vol. 44, no. 1, pp. 1–12, 2018.
- [17] P. A. Valverde, R. da Silva, and E. V. Stock, "Global oscillations in the optional public goods game under spatial diffusion," *Physica A: Statistical Mechanics and Its Applications*, vol. 474, no. 5, pp. 61–69, 2017.
- [18] L. Molleman, A. E. Quiñones, and F. J. Weissing, "Cultural evolution of cooperation: the interplay between forms of social learning and group selection," *Evolution and Human Behavior*, vol. 34, no. 5, pp. 342–349, 2013.
- [19] L. Shi, I. Romi, Y. Ma et al., "Freedom of choice adds value to public goods," *Proceedings of the National Academy of Sciences*, vol. 117, no. 30, Article ID 201921806, 2020.
- [20] C. Hauert, "Volunteering as red queen mechanism for cooperation in public goods games," *Science*, vol. 296, no. 5570, pp. 1129–1132, 2002.
- [21] M. Perc, J. J. Jordan, D. G. Rand, Z. Wang, S. Boccaletti, and A. Szolnoki, "Statistical physics of human cooperation," *Physics Reports*, vol. 687, no. 5, pp. 1–51, 2017.
- [22] J. Quan, W. Liu, Y. Chu et al., "Stochastic evolutionary voluntary public goods game with punishment in a Quasi-birth-and-death process," *Scientific Reports*, vol. 7, no. 1, Article ID 16110, 2017.
- [23] X.-j. Wang, R. Dong, and L. Chen, "The optimal control for promoting the cooperation in evolution game generated by Prisoner's Dilemma," *Acta Mathematica Scientia*, vol. 38, no. 1, pp. 73–92, 2018.
- [24] E. Pennisi, "How did cooperative behavior evolve?" *Science*, vol. 309, no. 5731, p. 93, 2005.
- [25] Z. Wang, M. Jusup, L. Shi et al., "Exploiting a cognitive bias promotes cooperation in social dilemma experiments," *Nature Communications*, vol. 9, p. 2954, 2018.
- [26] A. Mao, L. Dworkin, S. Suri et al., "Resilient cooperators stabilize long-run cooperation in the finitely repeated Prisoner's Dilemma," *Nature Communications*, vol. 8, p. 13800, 2017.
- [27] H.-T. Normann and B. Wallace, "The impact of the termination rule on cooperation in a prisoner's dilemma experiment," *International Journal of Game Theory*, vol. 41, no. 3, pp. 707–718, 2012.
- [28] R. L. Riolo, M. D. Cohen, and R. Axelrod, "Evolution of cooperation without reciprocity," *Nature*, vol. 414, no. 6862, pp. 441–443, 2001.
- [29] V. A. A. Jansen and M. van Baalen, "Altruism through beard chromodynamics," *Nature*, vol. 440, no. 7084, pp. 663–666, 2006.
- [30] Z. Tian, X. Gao, S. Su, et al., J. Qiu, "Vcash: a novel reputation framework for identifying denial of traffic service in internet of connected vehicles," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 3901–3909, 2020.
- [31] Z. Tian, X. Gao, S. Su, J. Qiu, X. Du, and M. Guizani, "Evaluating reputation management schemes of internet of vehicles based on evolutionary game theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5971–5980, 2019.
- [32] N. Gruenheit, K. Parkinson, B. Stewart et al., "A polychromatic 'green beard' locus determines patterns of cooperation in a social amoeba," *Nature Communications*, vol. 8, p. 14171, 2017.
- [33] J. Joshi and V. Guttal, "Demographic noise and cost of greenbeard can facilitate greenbeard cooperation," *Evolution*, vol. 72, no. 12, pp. 2595–2607, 2018.
- [34] E. S. Danka, E. C. Garcia, and P. A. Cotter, "Are CDI systems multicolored, facultative, helping greenbeards?" *Trends in Microbiology*, vol. 25, no. 5, pp. 391–401, 2017.
- [35] E. Cohen and D. Haun, "The development of tag-based cooperation via a socially acquired trait," *Evolution and Human Behavior*, vol. 34, no. 3, pp. 230–235, 2013.
- [36] D. K. Aanen and T. Czárán, "The early evolution of cooperation in humans. On cheating, group identity and group size," *Behaviour*, vol. 153, no. 9, pp. 1247–1266, 2016.
- [37] C. Taylor and M. A. Nowak, "Evolutionary game dynamics with non-uniform interaction rates," *Theoretical Population Biology*, vol. 69, no. 3, pp. 243–252, 2006.
- [38] R. Dong, X. J. Wang, and L. Chen, "The controller design for promoting the evolution of cooperation in the prisoner's dilemma based on thenon-uniforminteractionrates," *Systems Engineering Theory & Practice*, vol. 37, no. 10, pp. 2582–2591, 2017.
- [39] Y. Yuan, L. Guo, and Z. Wang, "Composite control of linear quadratic games in delta domain with disturbance observers," *Journal of the Franklin Institute*, vol. 354, no. 4, pp. 1673–1695, 2017.
- [40] T. L. Vincent and T. L. S. Vincent, "Evolution and control system design the evolutionary game," *IEEE Control Systems*, vol. 20, no. 5, pp. 20–35, 2000.
- [41] L. Wang and J. Du, "Evolutionary game theoretic approach to coordinated control of multi-agent systems," *Journal of Systems Science and Mathematical Sciences*, vol. 36, no. 3, pp. 302–318, 2016.
- [42] W. Weibull, *Evolutionary Game Theory*, MIT Press, Cambridge, UK, 1995.
- [43] D. Friedman, "On Economic applications of evolutionary game theory," *Journal of Evolutionary Economics*, vol. 8, no. 1, pp. 15–43, 1998.
- [44] M. A. Nowak, "Five rules for the evolution of cooperation," *Science*, vol. 314, no. 5805, pp. 1560–1563, 2006.
- [45] G.-Y. Tang, R. Dong, and H.-W. Gao, "Optimal sliding mode control for nonlinear systems with time-delay," *Nonlinear Analysis: Hybrid Systems*, vol. 2, no. 3, pp. 891–899, 2008.
- [46] Z.-M. Wang, A. Wei, X. Zhao, J. Yang, and G. Zong, "Stabilisation and \mathcal{H}_∞ control for switched port-controlled Hamiltonian systems with unstable modes and actuator

- saturation,” *International Journal of Systems Science*, vol. 51, no. 1, pp. 1–19, 2020.
- [47] Q. Liu and L. Long, “State-dependent switching law design with guaranteed dwell time for switched nonlinear systems,” *International Journal of Robust and Nonlinear Control*, vol. 30, no. 8, pp. 3314–3331, 2020.
- [48] Z. Song and J. Zhai, “Practical output tracking control for switched nonlinear systems: a dynamic gain based approach,” *Nonlinear Analysis: Hybrid Systems*, vol. 30, no. 11, pp. 147–162, 2018.
- [49] A. Basu, S. R. Bhaskaran, and R. Mukherjee, “An analysis of search and authentication strategies for online matching platforms,” *Management Science*, vol. 65, no. 5, pp. 2412–2431, 2019.