Forgiveness, Cooperation, and Present Bias in the Infinitely Iterated Prisoner's Dilemma

Minwook Kang
Department of Economics, Korea University

January 17, 2025

Abstract

How does present bias affect the propensity to forgive or punish? To explore this question, we employ quasi-hyperbolic discounting within the iterated Prisoner's Dilemma. The primary results from theoretical and simulation studies demonstrate that present bias impedes the sustainability of forgiving or lenient strategies (e.g., Tit-for-Tat with or without additional forgiveness, and Win-Stay, Lose-Shift) in a Nash equilibrium. However, this bias does not significantly impact the likelihood of sustaining cooperation for punitive strategies (e.g., the grim trigger).

Keywords: Present Bias; Tit-for-tat; Forgiveness Strategies; Iterated Prisoner's Dilemma; Cooperation.

JEL codes: C73; D64; D91.

.

1. Introduction

In the intricate world of self-interest and cooperation, the iterated prisoner's dilemma game provides a captivating glimpse into the realm of altruistic strategies. Since Axelrod's pioneering experiments in 1980, this game has sparked a flurry of research in economics, biology, and other fields.[1] This paper intends to explore the iterated prisoner's dilemma in the framework of the quasi-hyperbolic discounting model, as popularized by Laibson (1997), to comprehend the influence of human present bias pertaining to cooperation and forgiveness. Against the backdrop of empirical revelations, the present bias of human behavior emerges as a key player. Academics such as Thaler (1981) and Loewenstein (1987) have shed light on this temporal quirk, inspiring research outside the realm of exponential discounting, the standard practice in the iterated prisoner's dilemma.[2]

Due to the prevalence of exponential discounting in modeling the iterated prisoner's dilemma, it is essential to compare two distinct discounting models. To achieve this, we evaluate these discounters using a commonly employed measure of overall patience, which quantifies the area beneath the discounting functions.[3] By utilizing this measure, this paper examines whether mutual cooperation can be sustained in Nash equilibrium (NE) under various strategies: tit-for-tat with and without forgiveness (i.e., generous tit-for tat[4]), Win-Stay Lose-Shift (WSLS) and the grim trigger technique.[5] In the experiments by Dal Bó and Fréchette (2019), the main strategies observed were "tit-for-tat" and "grim trigger" (focus of this paper), along with "always defect." When real-world subjects participated in iterated prisoner's dilemma scenarios with perfect monitoring, these strategies predominated.[6]

[1] The results of a competition among fifteen programs developed for playing the iterated two-person Prisoner's Dilemma are detailed in Axelrod's (1980) work. In this tournament, the strategy 'Tit for Tat,' proposed by Anatol Rapoport, emerged as the champion. Axelrod and Hamilton (1981) along with Schofield (1984) delve into theoretical analyses of the outcomes.

[2] Human behavior deviates from a consistent time orientation, as evidenced by empirical and experimental findings (Thaler 1981, Loewenstein 1987, Benzion, Rapoport, and Yagil 1989, Kirby 1997, and Benhabib, Bisin, and Schotter 2010).

[3] Barro (1990) shows that under overall patience, there is no difference in the equilibrium allocations between quasi-hyperbolic and exponential economies. Specifically, only with observable data such as output, consumption, and savings, is it impossible to determine whether consumers make decisions based on quasi-hyperbolic or exponential discounting. After Barro (1990), the concept of overall patience has been widely used in the literature.

[4] An external chance of forgiveness, denoted as $q \in [0, 1)$, exists in the Generous Tit-for-Tat strategy by Nowak and Sigmund (1992). This forgiveness factor can improve long-term outcomes by facilitating cooperation recovery and curbing endless retaliation.

[5] In the theoretical model, this paper focuses only on the case where players aim to maintain mutual cooperation in a NE. However, in the simulation, we allow for the possibility of a non-cooperative equilibrium returning to a mutual cooperation equilibrium.

[6] However, the "always defect" strategy, which involves constant betrayal, cannot be sustained in a NE. This is precisely why we are not interested in this strategy.

Under the grim trigger strategy, the NE conditions for cooperation in both discounting models are identical.[7] However, this paper demonstrates that employing forgiving or lenient strategies such as tit-for-tat or WSLS requires a higher level of patience for quasi-hyperbolic discounters than for exponential discounters to maintain cooperation as a Nash Equilibrium (NE). This suggests that achieving a cooperative equilibrium becomes more difficult when present bias is high.

The primary finding of this study is that present bias makes sustaining forgiveness strategies more difficult than sustaining punishment strategies. Forgiveness primarily serves to motivate adversaries to cooperate and achieve mutually beneficial present and future payoffs. However, the presence of present bias reduces the propensity of players to prioritize present payoffs. Consequently, the tendency for players to deviate from cooperative strategies to secure higher present-value guarantees is greater, even if they are aware that this may have negative effects on their future payoffs.

In the theoretical approach, we focus on how the present bias affects the sustainability of mutual cooperation in a NE. These theoretical results provide a clear understanding of how NEs are formed in a deterministic world with quasi-hyperbolic discounting preferences. However, uncertainty in the real world may cause them to modify their decisions in response to the altered environment. This dynamic situation is simulated in this paper. In the simulation, payoffs are treated as random variables based on an AR(1) model, and forward-looking players choose whether to cooperate or defect in each period. We iterate the game a sufficient number of times to observe the frequency with which the mutual cooperation state (or tit-for-tat state) is maintained. The simulation reveals several important findings. In accordance with our theoretical predictions, present bias negatively affects mutual cooperation, resulting in a decrease in social welfare. The negative welfare effects caused by present bias are exacerbated when players employ forgiveness-oriented strategies, indicating that the adoption of such strategies in societies with highly present-biased individuals becomes challenging. In addition, the simulation demonstrates that there exists an optimal level of forgiveness that maximizes social welfare for any given level of present bias.

Prior research has incorporated non-geometric discounting, such as the quasi-hyperbolic type, into repeated games.[8] For instance, Chade et al. (2008) used quasi-hyperbolic discounting to explain equilibrium payoffs in subgame perfect equilibrium, revealing non-monotonic

[7]Musau (2013) also compared two discounters employing a grim trigger technique. In contrast, the exponential discounter consistently demonstrates greater overall patience than the quasi-hyperbolic discounter in his comparison. In his setup, the discounting factor for the exponential discounter is identical to the long-term discounting factor (denoted as δ) for the quasi-hyperbolic discounter, resulting in this disparity.

[8]There are a broad range of non-geometric discounting methods, including myopic and future-biased discounting. Quasi-hyperbolic discounting is one of the main forms of myopic discounting (Kang and Wang 2019 and Kang 2020).

behavior regarding a hyperbolic discount factor.[9] Obara and Park (2017) found a variety of worst punishment equilibria across various bias categories. The worst punishment for future-biased players resembles a combination of rewards and penalties, whereas for present-biased players, it fluctuates. Meanwhile, Kim (2023) conducted experiments in which participants engaged in multi-round games with rewards distributed gradually (weekly/monthly). Adjusting rewards influenced discounting, and delaying first-round rewards affected present bias in quasi-hyperbolic discounting. The fact that weekly rewards boosted cooperation indicates that present bias hinders it. In contrast to antecedent studies, this paper focuses on specific strategies, such as grim trigger and generous tit-for-tat, which incorporate uncertainties. These elements can be linked to simulations and provide intuitive insights into how present bias modulates behaviors associated with forgiveness and punishment.

Hyperbolic discounting has been mostly applied to multiple-period consumption-savings models where each period consumer makes decisions on savings rates on a continuous domain. The present bias in each future period affects the savings rates in those periods, thus the current self needs to consider all the dynamic inconsistency in the decision. On the other hand, in the prisoner's dilemma game, the decisions are discretely simplified as cooperation or defect. Therefore, under Nash equilibrium, there is no incentive for future selves to deviate from cooperation, even under a quasi-hyperbolic discounting function. Hyperbolic discounting affects decision making in two ways: One is the direct bias effect such that the change of the discounting function affects decisions in each period. The other effect is the dynamic inconsistent effects where the inconsistency of future and current preferences makes the current self strategically make decisions. The approach of the Nash equilibrium in this paper naturally focuses on the first effect.[10]

The remainder of this paper is structured as follows. Section 2 presents the primary model outlined in this paper. Section 3 analyzes how present bias affects the effectiveness of the tit-for-tat and generous tit-for-tat strategies in the context of the iterated prisoner's dilemma. Section 4 explores the impact of present bias on equilibrium when players employ forgiveness-oriented strategies. Sections 5 and 6 investigate the implications of present bias

[9] Chade et al. (2008) show that when the minimax point of the stage game is a Nash equilibrium, as in the Prisoner's Dilemma, the worst equilibrium and continuation payoffs coincide, restoring monotonicity. Therefore, in the Prisoner's Dilemma setting, there is no non-monotonic relation in the β-δ model. In this paper, the cooperation rate is non-monotonically concave in the degree of forgiveness, a property known from Axelrod (1980) under exponential discounting.

[10] Specifically, in the consumption-savings model, assume that consumers make decisions based on subgame perfect Nash equilibrium in a three-period model. Savings in periods 1 and 2 are s_1 and s_1. The period-2 savings function, $s_2(s_1)$, depends on s_1. The equilibrium $(s_1,\ s_2(s_1))$ shows how period-1 savings s_1 influences s_2, reflecting time inconsistency. On the other hand, this paper examines mutual cooperation in a Nash equilibrium. Once achieved, players choose 'cooperate' for all periods, meaning future decisions are unaffected by previous ones, which implies that there is no time-inconsistency issue.

on the WSLS and grim trigger strategies, respectively. Section 7 presents the simulations in which nondeterministic payoffs for mutual cooperation are assumed. Finally, Section 8 concludes this study.

2. The Model

There are two players, and each player has the option of choosing between two actions: Cooperate or Defect. If both players choose to cooperate, each obtains a payoff of $C \in \mathbb{R}$. If one player chooses defect and the other chooses cooperate, the former obtains a payoff of $A \in \mathbb{R}$ and the latter obtains a payoff of $Z \in \mathbb{R}$. If both players choose defect, each obtains a payoff of $D \in \mathbb{R}$. Following typical iterated prisoner's dilemma game, we assume that $A > C > D > Z$ and $C > (A + Z)/2$. The payoff matrix should be

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	C, C	Z, A
	Defect	A, Z	D, D

The iterated prisoner's dilemma game is a triplet: $G = (N, S, u)$, where $N = \{1, 2\}$ is the set of players; $S = \times_{i \in N} S^i$ is the set of strategy profile, such that $S^1 = S^2 = \{$cooperate, defect$\}$; and u is a payoff function such that $u^i : S \to \mathbb{R}$. U^i is the discounted payoff function in period t, such that $U : S \to \mathbb{R}^N$. The discounted payoff function of player i at period t is defined by

$$U^i = \sum\nolimits_{t=0}^{\infty} D(t) u^i(s(t)).$$

where $s(t)$ is the stage-game outcome played in period t and $D(t)$ is a discount function. Each participant selects the action that maximizes the discounted payoff. A discrete discount function, $D(t)$, of the exponential discounting is given by

$$D^E(t) = \gamma^t, \quad t = 0, 1, 2, \cdots ,$$

where $\gamma \in (0, 1)$ is the discount rate and $t \in \{0, 1, 2...\}$ represents a time delay. Under the quasi-hyperbolic discounting factor of $(\beta, \delta) \in (0, 1)^2$, the discounting function is given by

$$D^H(t) = \begin{cases} 1 & \text{if} \quad t = 0 \\ \beta\delta^t & \text{if} \quad t = 1, 2, ... \end{cases} .$$

For the comparison between less present bias (including exponential discounting) and more present bias, we employ a commonly used measure of overall patience, which computes the area below the discounting function (e.g., Myerson et al. 2001, Chade et al. 2008, Caliendo and Findley 2014, Strulik 2015, Cabo et al. 2015, Kang 2021; 2025, Kang and Kim 2024). We define the overall measure of patience, $P(D(t))$ as

$$P(D(t)) = \sum\nolimits_{t=0}^{\infty} D(t), \tag{1}$$

Fixing overall levels of impatience identical, from Eq. (1) we obtain the following results:

$$\frac{1}{1-\gamma} = \frac{1-\delta+\beta\delta}{1-\delta}, \tag{2}$$

In Eq. (2), if $\beta = 1$ (no present bias), we have $\gamma = \delta$. For the same level of overall patience, we compare any two discounters, one of which is less present bias (denoted as (β^-, δ^-)), and the other is more present bias (denoted as (β^+, δ^+)), where $\beta^+ < \beta^-$. Under equal patience (i.e., $(1-\delta^- + \beta^-\delta^-)/(1-\delta^-) = (1-\delta^+ + \beta^+\delta^+)/(1-\delta^+)$), $\beta^+ < \beta^-$ implies $\delta^+ > \delta^-$.

3. Present Bias and Cooperation under Tit-for-Tat

Tit-for-tat is a well-studied game theory strategy characterized by reciprocal cooperation and defection. Its simple yet effective nature has sparked widespread interest, particularly in Axelrod's seminal work "The Evolution of Cooperation" (1984). This demonstrates how tit-for-tat promotes stability and cooperation in iterated prisoner's dilemma games. Further insights from Nowak and Sigmund's "Generous Tit-for-tat" (1992) explored cooperation in complex social systems. This section investigates the effect of present bias on the effectiveness of tit-for-tat and generous tit-for-tat in the iterated prisoner's dilemma.

The following strategy describes a (generous) tit-for-tat strategy. Player $i \in N$ is said to be playing a tit-for-tat strategy with forgiveness rate $q \in [0, 1)$ if for every period $\tau = 1, 2, \cdots$,

$$s_\tau^i = \begin{cases} \text{Cooperate} & \text{if } \; s_{\tau-1}^j = \text{Cooperate} \\ \text{Cooperate with probability } q & \\ \text{Defect with probability } 1-q & \text{if } \; s_{\tau-1}^j = \text{Defect} \end{cases} \tag{3}$$

If $q = 0$, the strategy in Eq. (3) is identical to standard tit-for-tat without forgiveness. The strategy with $q = 100\%$ results in always-cooperation, which is not supported by NE. The results of Anatol Rapoport's experiment indicate subsequent analysis of Axelrod's Experiments, revealing that incorporating a small probability of forgiveness (e.g., $q = 1\text{˜}5\%$) into tit-for-tat strategies yields a superior outcome compared to abstaining from forgiveness entirely (i.e., $q = 0$). It is well-known that, in general, a subgame perfect equilibrium with the tit-for-tat strategy does not exist. Nonetheless, a NE exists if the discount factors are sufficiently large. This paper assumes that q is sufficiently low and γ (or (β, δ)) is sufficiently large to ensure the existence of the NE.

Under quasi-hyperbolic discounting (and exponential discounting if $\beta = 1$), the condition for cooperation to be sustained in a NE under tit-for-tat strategy is

$$C + \frac{\beta\delta}{1-\delta}C \geq A + \beta\delta V(q, \delta), \tag{4}$$

where

$$\begin{aligned} V(q,\delta) &= \underbrace{q}_{\text{opponent's forgiveness}} \times \left(\frac{C}{1-\delta}\right) \\ &+ \underbrace{(1-q)}_{\text{opponent's punishment}} \times \\ &\left(Z + \delta\left(\underbrace{q}_{\text{agent's forgiveness}} \frac{C}{1-\delta} + \underbrace{(1-q)}_{\text{agent's punishment}} (A + \delta V(q,\delta))\right)\right) \end{aligned} \tag{5}$$

With a positive value of $q > 0$, we must solve the discounted payoff recursively as expressed in Eqs. (4-5). In the inequality in Eq. (4), $C + \frac{\beta\delta}{1-\delta}C$ represents the agent's discounted payoff where mutual cooperation is sustained in the infinite-period game. Meanwhile, $A + \beta\delta V(q, \delta)$ in Eq. (4) represents the discounted payoff in the case where the player chooses to defect. At the time of choosing defect, the payoff should be A. However, in the subsequent period, the opponent chooses to work with probability q (forgiveness) or to defect with probability $1 - q$ (punishment). In the event that the opponent overlooks the agent's flaw, the status reverts to one of mutual cooperation. However, with punishment, the agent's payoff should be Z in the next period, and the agent will cooperate with probability q or defect with probability $1 - q$ in the subsequent period.

With high value of q, the NE does not exist. Specifically, if $q = 1$ (always forgive), the $A + \beta\delta V(q, \delta)$ is equal to $A + \frac{\beta\delta}{1-\delta}C$, which is strictly higher than $C + \frac{\beta\delta}{1-\delta}C$. Therefore, our analysis focuses on the tit-for-tat ($q = 0$) and a small level of forgiveness. For the same level of overall patience, the left-hand side in Eq. (4) remains invariant with present bias, i.e., $C + \frac{\beta^-\delta^-}{1-\delta^-}C = C + \frac{\beta^+\delta^+}{1-\delta^+}C$ (since $\frac{\beta^-\delta^-}{1-\delta^-} = \frac{\beta^+\delta^+}{1-\delta^+}$). Therefore, the impact of the present bias on the sustainability of cooperation is determined by the right-hand side in Eq. (4). The following Proposition shows that for a small value of $q \geq 0$, present bias still makes the cooperation to be sustained.

Proposition 1 *(Tit-for-tat with/without forgiveness): There always exists a value* $\overline{q}$*, such that for any* $q \in [0, \overline{q})$ *if the cooperation equilibrium is sustained in a NE in the game with* (β^+, δ^+)*, cooperation is also sustained in the game with* (β^-, δ^-)*. However, the reverse is not necessarily true.*

Proposition 1 implies the difficulty in sustaining mutual cooperation in NE under both the standard and generous tit-for-tat strategies due to the influence of present bias. The simple intuition underlying this result is that under present bias, agents place a greater emphasis on immediate rewards and a lesser emphasis on future rewards. When an agent deviates from cooperation, they are eligible for a greater immediate reward, $A(> C)$. Due to present bias, the immediate large reward of A has a greater impact on the discounted payoff, resulting in difficulties sustaining mutual cooperation. In the following section, we investigate how present bias inhibits the adoption of a generous forgiveness strategy in the tit-for-tat strategy.

4. Present bias and Forgiveness

The preceding section demonstrates the adverse impact of present bias on cooperation when (generous) tit-for-tat strategies are utilized. Subsequently, it should be determined how present bias affects the equilibrium when players adopt forgiveness-oriented strategies. The primary conclusion of this section is that dealing with present bias suggests difficulties in maintaining forgiveness tactics versus punishment tactics. In particular, the following proposition demonstrates that with a greater present bias (i.e., a lower value of β), an increase in the forgiveness parameter, q, makes it more difficult to sustain cooperation within a NE:

Proposition 2 *There always exists* $\overline{q} > 0$*, such that for any* $q^l, q^h \in [0, \overline{q})$ *and* $q^l < q^h$*,*

$$\beta^-\delta^-V(q^h, \delta^-) - \beta^-\delta^-V(q^l, \delta^-) < \beta^+\delta^+V(q^h, \delta^+) - \beta^+\delta^+V(q^l, \delta^+) \tag{6}$$

The change in forgiveness level (q) does not impact the discounted payoff under cooperation; however, it has an effect on the discounted payoff with deviation. According to Proposition 2, a stronger present bias (i.e., under (β^+,δ^+)) amplifies the additional gain in discounted payoff with deviation, as the forgiveness level increases from q^l to q^h. This suggests that the present bias discourages players from adopting a high-level forgiving strategy in repeated games.

Defining $\overline{C}$ as the threshold (minimum) value of C for the sustainability of cooperation, we can comprehend the inequality better in Eq. (6). Recalling the condition for cooperation to be sustained in NE, we have

$$C + \frac{\beta\delta}{1-\delta}C \geq A + \beta\delta V(q,\delta).$$

Since the left-hand side is increasing in C, we define the threshold value $\overline{C}$ such as

$$\overline{C} + \frac{\beta\delta}{1-\delta}\overline{C} = A + \beta\delta V(q,\delta),$$

where the necessary and sufficient conditions for the sustainability of cooperation in the game is $C > \overline{C}$. Then, for the two different values of forgiveness levels q^l and q^h, where $q^l < q^h$, we obtain

$$\overline{C}^h + \frac{\beta\delta}{1-\delta}\overline{C}^h = A + \beta\delta V(q^h,\delta) \tag{7}$$

and

$$\overline{C}^l + \frac{\beta\delta}{1-\delta}\overline{C}^l = A + \beta\delta V(q^l,\delta). \tag{8}$$

From Eqs. (7-8), we have

$$\left(\overline{C}^h - \overline{C}^l\right) \underbrace{\frac{1-\delta+\beta\delta}{1-\delta}}_{\text{overall-patience (fixed)}} = \underbrace{\left\{\beta\delta V(q^h,\delta) - \beta\delta V(q^l,\delta)\right\}}_{\text{decreases with increased } \beta \text{ by Eq. (25)}} \tag{9}$$

In Eq. (9), $\frac{1-\delta+\beta\delta}{1-\delta}$ is the fixed level of overall patience. The term $\beta\delta V(q^h,\delta) - \beta\delta V(q^l,\delta)$ in Eq. (9) increases with decreased β (or increased level of present bias). Therefore, we know that $\overline{C}^h - \overline{C}^l$ is higher when present bias is stronger. This suggests that present bias and forgiveness are complementary factors that negatively affect the cooperation's sustainability. From Eq. (9), we obtain the following proposition:[11]

Proposition 3 *There always exists $\overline{q} > 0$, such that for any $q \in [0, \overline{q})$, $d\overline{C}/dq$ increases with increased present bias (i.e., with decreased β) where $\overline{C}$ represents the threshold value for cooperation to be sustained in a NE.*

For better understanding of Proposition 3 and Eq. (9), we calculate the threshold value ($\overline{C}$) using the payoff matrix in Dal Bó and Fréchette (2019) as A=50, D=25, and Z=12.

[11] This approach of deriving a threshold value is similar to classic research on risk aversion, where the impact of risk aversion is easier to understand based on consumption goods using the concept of a certainty equivalent, rather than utility values. The threshold value also facilitates the design of simulations.

Setting $\gamma = 0.5$, we demonstrate in Table 1 the impact of forgiveness on the value of $\overline{C}$ under different values of β and q.

	$\beta = 1$		$\beta = 0.7$		$\beta = 0.5$	
	$\overline{C}$	$\overline{C}^h - \overline{C}^l$	$\overline{C}$	$\overline{C}^h - \overline{C}^l$	$\overline{C}$	$\overline{C}^h - \overline{C}^l$
$q = 0\%$	37.33		38.04		38.60	
$q = 10\%$	38.21	0.87	39.05	1.02	39.82	1.22
$q = 20\%$	39.14	0.94	40.09	1.04	40.99	1.17
$q = 30\%$	40.15	1.01	41.15	1.059	42.11	1.13

Table 1: The impact of forgiveness (q) on the value of $\overline{C}$ under different β values

In Table 1, as forgiveness increases from 0% to 10%, $\overline{C}$ increases by 0.87 for β=1, 1.02 for β=0.7, and 1.22 for β=0.5 ($0.87 < 1.02 < 1.22$). These patterns of increase are also observed when q increases from 10% to 20% and from 20% to 30%, as shown in Proposition 3. With a stronger present bias, the generous forgiveness strategy requires a greater payoff from mutual cooperation to remain viable.

5. Win-Stay, Lose-Shift

The Win-Stay, Lose-Shift (WSLS) strategy is also a well-studied approach in the Prisoner's Dilemma game and has been proven effective in the literature, for example, by Nowak and Sigmund (1993) and Imhof et al. (2007). WSLS repeats the previous choice if the payoff is higher than a satisfactory level and changes otherwise. In the Prisoner's Dilemma, the payoffs from (s_i, s_j)=(Cooperate, Cooperate) and (s_i, s_j)=(Defect, Cooperate) are satisfactory for player i, while (s_i, s_j)=(Cooperate, Defect) and (s_i, s_j)=(Defect, Defect) are not. Therefore, player i cooperates if the previous game was (s_i, s_j)=(Cooperate, Cooperate) or (s_i, s_j)=(Defect, Defect) and defects if the previous game was (s_i, s_j)=(Defect, Cooperate) or (s_i, s_j)=(Cooperate, Defect).

The WSLS strategy is considered forgiving or lenient because it allows players to return to cooperative behavior after a defection. Specifically, if two players defect in the previous round, resulting in losses for both, they will then cooperate in the next round. This characteristic allows it to reset back to cooperative behavior after deviations, which is a key aspect of forgiveness in repeated game strategies. This makes WSLS less strict compared to strategies like the grim trigger, which permanently punishes any defection by never returning to cooperation.

The WSLS strategy is

$$s_\tau^i = \begin{cases} \text{Cooperate} \quad \text{if} \quad \left(s_{\tau-1}^i, s_{\tau-1}^j\right) = \begin{array}{c} \text{(Cooperate, Cooperate)} \\ \text{or} \\ \text{(Defect, Defect)} \end{array} \\ \\ \text{Defect} \quad \text{if} \quad \left(s_{\tau-1}^i, s_{\tau-1}^j\right) = \begin{array}{c} \text{(Defect, Cooperate)} \\ \text{or} \\ \text{(Cooperate, Defect)} \end{array} \end{cases} \tag{10}$$

Under the WSLS strategy, consider the situation where one of the players deviates from mutual cooperation. In the next game, both players choose to defect based on the strategy in Eq. (10). In the subsequent game, both players choose to cooperate based on Eq. (10), and consequently, mutual cooperation is recovered. Therefore, the condition for cooperation to be sustained in a Nash Equilibrium (NE) under the WSLS strategy is

$$C + \frac{\beta\delta}{1-\delta} C \geq A + \beta\delta \left(D + \delta C + \delta^2 C + \delta^3 C + \cdots\right), \tag{11}$$

Proposition 4 *If the cooperation equilibrium is sustained in a WSLS in the game with (β^+, δ^+), cooperation is also sustained in the game with (β^-, δ^-). However, the reverse is not necessarily true.*

Proposition 4 suggests that present bias makes it difficult to maintain mutual cooperation in Nash Equilibria (NE) under Win-Stay, Lose-Shift (WSLS) strategies. When an agent opts out of cooperating, they receive a larger immediate benefit, denoted by A (which is greater than D). The influence of present bias amplifies the impact of this immediate large reward of A on the discounted payoff, thereby complicating the sustainability of mutual cooperation.

6. Grim trigger

A player employing the grim trigger strategy will initially cooperate. However, as soon as the opponent defects (meeting the trigger condition), the player using grim trigger will also defect for the remainder of the game. Player $i \in N$ is said to be playing a trigger strategy if for every period τ where $\tau = 1, 2, ...,$

$$s_\tau^i = \begin{cases} \text{Cooperate} & \text{if} \quad s_\tau^i = s_\tau^j = \text{cooperate for all } t = 1, 2, ..., \tau - 1 \\ \text{Defect} & \text{if} \quad \text{otherwise.} \end{cases} \tag{12}$$

With a trigger strategy, present bias has no effect on the sustainability of mutual cooperation in a NE.

Proposition 5 *Under grim trigger strategy, mutual cooperation is sustained in a NE with* (β^-, δ^-)*, if and only if the cooperation is sustained in the game with* (β^+, δ^+)*.*

In cases where there is no forgiveness, such as the grim trigger strategy, Proposition 5 shows that the degree of present bias has no effect on the NE's sustainability. Conversely, as demonstrated in Proposition 3, as the degree of forgiveness increases in the tit-for-tat strategy, it becomes more difficult to maintain mutual cooperation in equilibrium due to present bias. The results of the tit-for-tat and grim strategies indicate that present bias impedes the benefits of forgiveness-based strategies but has no significant impact on punishment-based strategies.

7. Simulation with dynamic decisions

In the iterated prisoner's dilemma, the results of the previous sections provide theoretical evidence that a present bias hinders both cooperation and forgiveness. The theoretical predictions would precisely depict the equilibrium, but in the absence of environmental uncertainty, the real world introduces payoff uncertainties. In such situations, the theoretical model would be limited in accurately predicting the results. In this section, we conduct simulations with the assumption that the payoffs from mutual cooperation are not deterministic. The two primary goals of these simulations are as follows: (1) to demonstrate the consistency between theoretical predictions and equilibrium consequences in a stochastic dynamics environment; and (2) to quantify the impact of present bias on the optimal level of forgiveness and the welfare of the players.

We consider a scenario in which both players employ the generous tit-for-tat tactic. In the absence of forgiveness, the equilibrium typically remains either in a state of mutual cooperation or the tit-for-tat. However, if players adopt possibility of forgiveness, the tit-for-tat states can transition to mutual cooperation. In addition, we assume that the payoff resulting from mutual cooperation is random rather than deterministic. If a player anticipates a low payoff from cooperation in future periods, they may choose to deviate from mutual cooperation. These two additional elements in the typical model make it possible to create a dynamic model in which players are not necessarily bound to a single state.

We use the following payoff matrix, including the following uncertainties:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$\widetilde{R}, 40$	$12, 50$
	Defect	$50, 12$	$25, 25$

Specifically, we assume that player 1's payoff in a state of mutual cooperation is uncertain.[12] We assume that $\widetilde{R}$ follows AR(1) model, such as

$$\widetilde{R}_t = R_{t-1} + s((40 + d) - R_{t-1})\varepsilon_t - s(R_{t-1} - (40 - d))\mu_t, \tag{13}$$

where ε_t and μ_t are independently and identically distributed random variables uniformly distributed in the interval [0,1]. Then, we have

$$E\left[\widetilde{R}_t\right] = 40,\ \widetilde{R}_t \in [40 - d, 40 + d]$$

The conditional expected value and variance of $\widetilde{R}_t$ given R_{t-1} is

$$E\left[\widetilde{R}_t | R_{t-1}\right] = (1 - s)R_{t-1} + s40 \tag{14}$$

and

$$Var\left[\widetilde{R}_t | R_{t-1}\right] = \frac{s^2\left(d^2 + (R_{t-1} - 40)^2\right)}{6} \tag{15}$$

Eqs. (13–15) indicate that with higher level of value $s \in (0, 1)$, $\widetilde{R}_t$ is more likely to deviated from R_{t-1}. Players are risk-neutral and employ a tit-for-tat strategy with forgiveness.

Each player makes decisions at the beginning of each period prior to the realization of player 1's payoff under mutual cooperation. Player 1 has an incentive to deviate from mutual cooperation if

$$C + \frac{\beta\delta}{1 - \delta}C < E\left[\widetilde{R}_t | R_{t-1}\right] + \beta\delta V(q, \delta). \tag{16}$$

The value of the right-hand side in Eq. (16) can be computed from the value of R_{t-1} using Eqs. (13) and (14). In the game, the states of mutual cooperation and tit-for-tat are randomly repeated as described in Figure 1.

[12] In this straightforward model, player 2's payoff is deterministic. If player 2's payoff becomes uncertain, player 1 should consider the possibility that player 2 will deviate from mutual cooperation, which may result in unnecessary mathematical complexities in determining decisions.

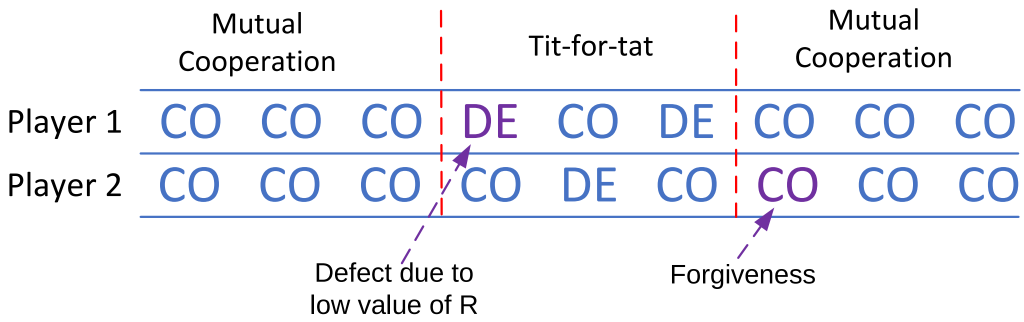

Figure 1: Randomly Repeated States of Mutual Cooperation and Tit-for-Tat, where CO and DE denote Cooperate and Defect, respectively.

We conduct simulations for the number of period $T = 1,000$ with different parameters of (β, q). Such simulations are conducted 900 times to derive the mean and 95% confidence interval for the cooperation rates. With a higher value of iteration, the mean value of payoffs and percentage of mutual cooperation states converge. In the simulation, we set $\gamma = 0.5$. If R settles at 40 with a certainty level and $q = 0$, all parameter choices of (β, δ) in the simulation guarantee the sustainability of cooperation in the Nash equilibrium. Specifically, the threshold of γ required to sustain cooperation as an equilibrium outcome is 0.357,[13] which is lower than 0.5. For low values of β, cooperation would not be sustained in a Nash equilibrium: the minimum value of β for cooperation to be sustained where $(R, q) = (40, 0)$ is $\beta = 0.11$ under equal patience with $\gamma = 0.5$.

Figure 2 plots how mutual-cooperation rates vary for different values of β and q with $(s, d) = (0.1, 8)$. The dashed curves represent 95% confidence intervals in the simulation, while the middle curve represents the mean values of the simulations. It shows that (1) the cooperation rate is higher with a higher value of β (i.e., weaker present bias), which could be implied from Proposition 1, and (2) the cooperation rate tends to be concave in the forgiveness level, indicating an optimal forgiveness level that maximizes the cooperation rate. If the forgiveness level is too low, the likelihood of a tit-for-tat state returning to mutual cooperation is also low. Conversely, if the level of forgiveness is too high, as demonstrated by Proposition 3, the player has a greater incentive to deviate from mutual cooperation when the expected value of the reward for mutual cooperation is low. Therefore, there is, in general, an interior value of forgiveness that maximizes the cooperation rate. According to the

[13]The threshold value of $\gamma(= 0.357)$ where $q = 0$ can be computed by the following inequality:

$$\frac{40}{1-\gamma} \geq \frac{50}{1-\gamma^2} + \frac{12\gamma}{1-\gamma^2}$$

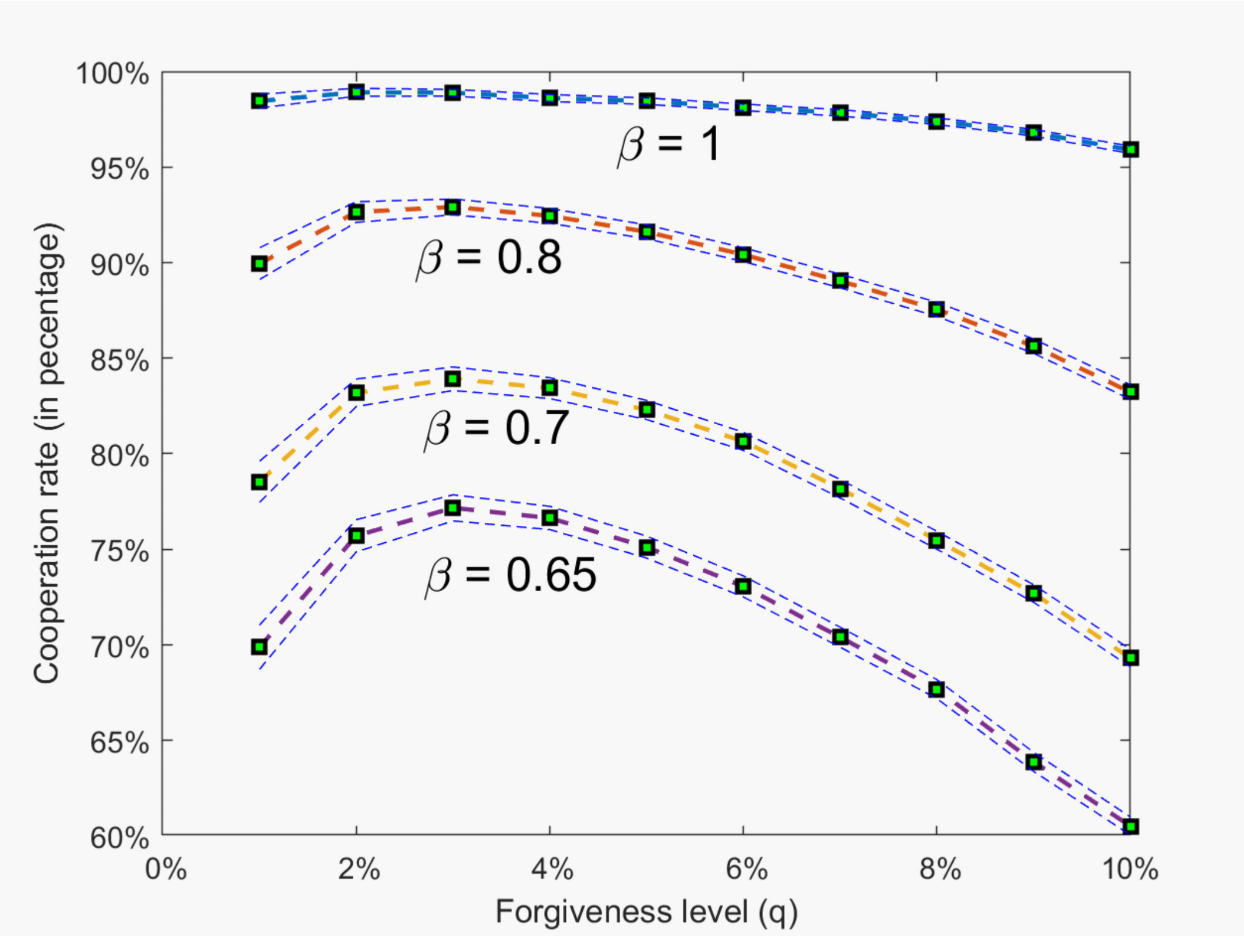

Figure 2: Mutual Cooperation rates for different values of (β,q) where (s,d) = (0.1,8).

characteristics of the iterated prisoner's dilemma, the rate of mutual cooperation and social welfare (the sum of the two players' payoffs) are so highly correlated that the specific result of social welfare is not specified in this paper. Lastly, the simulation results demonstrate that for lower values of β (i.e., stronger present bias), the reduction in the percentage of cooperation intensifies with a higher level of forgiveness, as shown in Proposition 3. Specifically, when β equals 1, the cooperation rate only diminishes by 2.6 percentage points (=98.6%-96.0%) when the forgiveness rate increases from 5% to 10%. However, for lower values of β (β=0.8, 0.7, 0.65), the cooperation rate experiences a higher decrease of 7.7, 12, and 14.9 percentage points, respectively, with the same change in forgiveness rate.

An increase in the values of (s, d), in general, lowers the cooperation rates as it causes the payoff to deviate further from 40, increasing the chance of the payoff being low enough to provide incentive for deviation from the cooperative state. Our simulation results unequivocally demonstrate that elevated values of either s or d, or a decreased value of γ, decrease the cooperation rates. This holds true regardless of the chosen parameters for (β, q). Ex-

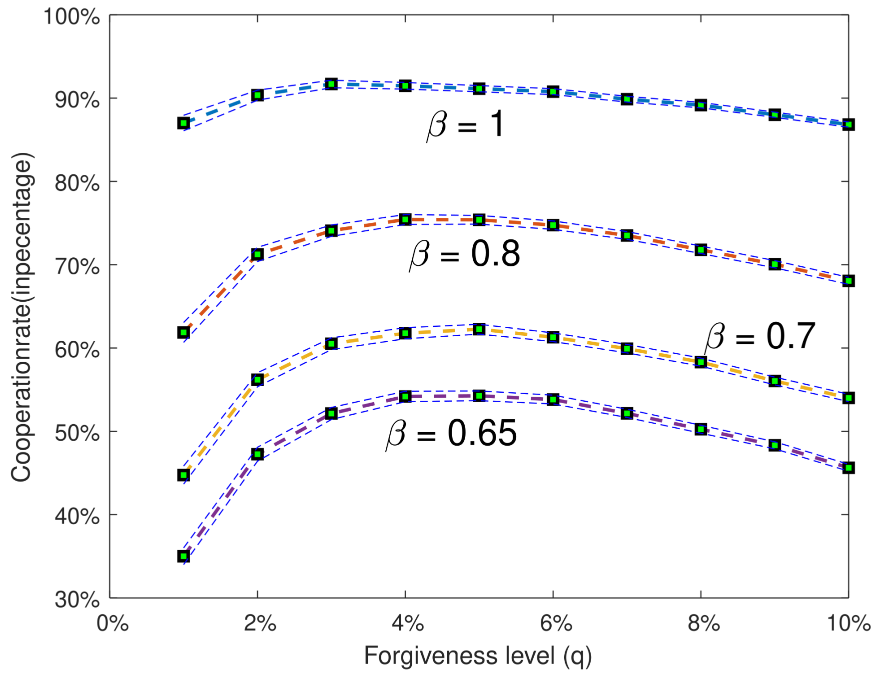

Figure 3: Mutual Cooperation rates for different values of (β,q) where (s,d) = (0.15,8).

amining the simulation outcomes for distinct sets of (s, d) values reveals that they exhibit similar patterns. These observed consistencies are consistent with the theoretical outcomes described in this article as shown in Figures 3 and 4.

8. Conclusion

The iterated prisoner's dilemma game has been instrumental in elucidating how self-interested individuals can adopt altruistic strategies, and has found applications in a variety of research fields. This paper focused on the iterated prisoner's dilemma within the framework of quasi-hyperbolic discounting, a model that effectively addresses present bias and time inconsistency in human behavior. We examined the impact of two discounting models, exponential and quasi-hyperbolic, on sustaining cooperation in NE under different strategies by comparing the discounting models.

Our findings indicate that the quasi-hyperbolic discounter requires a higher level of overall patience than the exponential discounter to maintain cooperation as a NE when employing the tit-for-tat and WSLS strategy. This indicates that sustaining cooperative equilibria becomes more difficult as current bias increases. Moreover, we have observed that present bias makes forgiveness strategies more difficult to sustain than punishment strategies. The

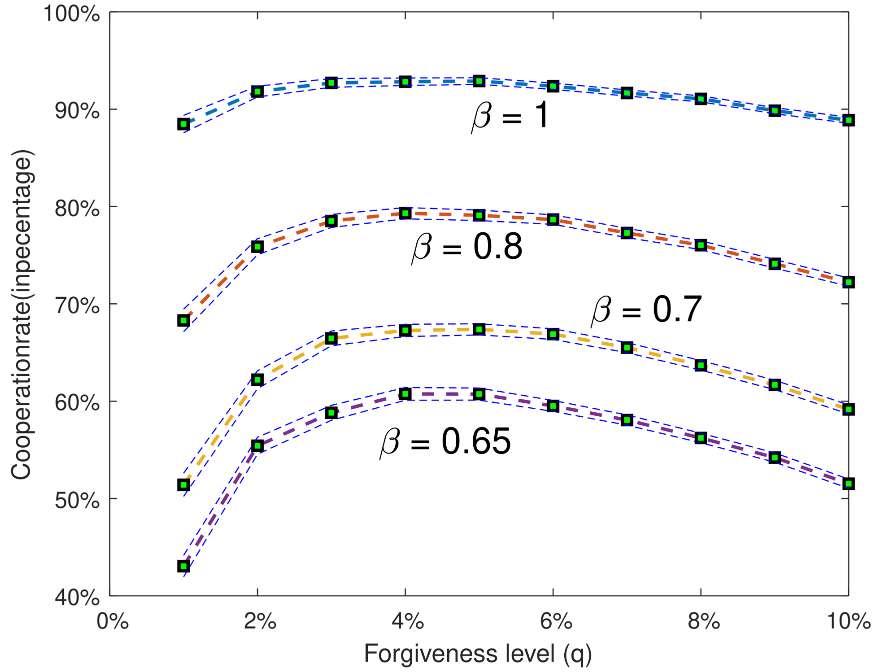

Figure 4: Mutual Cooperation rates for different values of (β,q) where (s,d) = (0.1,10).

propensity to prioritize immediate benefits reduces the inclination of players to engage in cooperative behavior, even if it has negative long-term consequences.

Our theoretical approach has provided us with valuable insights into the effects of present bias on the sustainability of mutual cooperation in a NE. However, we acknowledge that decisions in the real world are influenced by uncertainty and dynamic environments. To account for this, we conducted simulations of the dynamic situation using random variables as payoffs in a simple autoregressive model. The simulation's outcomes corroborate our theoretical findings, demonstrating that present bias has a negative effect on mutual cooperation, resulting in a decline in social welfare. In addition, the simulation demonstrates that there is an optimal level of forgiveness that maximizes social welfare, thereby providing a means of mitigating the negative effects of present bias.

In conclusion, our study sheds light on the difficulties posed by present bias in sustaining cooperative strategies and highlights the importance of forgiveness in incentivizing cooperation. Understanding the dynamics of quasi-hyperbolic discounting and its implications on decision-making can provide valuable insights for a variety of disciplines, from the social sciences to economics, and strategies for achieving long-term cooperation and welfare improvements.

Appendices

A. Proof of Proposition 1

For the proof of Proposition 1, we need the following lemma:

Lemma 1 *If $q = 0$, $A + \beta^- \delta^- V(0, \delta^-) < A + \beta^+ \delta^+ V(0, \delta^+)$.*

Proof. From Eq. (5), we can derive $V(q, \delta)$ as

$$V(q, \delta) = \frac{A\delta(1-\delta)(1-q)^2 + Cq\left(1 + \delta(1-q)\right) + Z(1-\delta)(1-q)}{(1-\delta)\left(1-\delta^2(1-q)^2\right)}, \tag{17}$$

If $q = 0$, from Eq. (17), we have

$$\beta\delta V(0, \delta) = \beta\delta\frac{A\delta(1-\delta) + Z(1-\delta)}{(1-\delta)\left(1-\delta^2\right)} = \frac{\beta\delta}{1-\delta}\frac{Z + A\delta}{1+\delta}. \tag{18}$$

Since $\frac{\beta^- \delta^-}{1-\delta^-} = \frac{\beta^+ \delta^+}{1-\delta^+}$, $Z < A$, and $\delta^+ > \delta^-$, from Eq. (18) we have

$$\beta^- \delta^- V(0, \delta^-) < \beta^+ \delta^+ V(0, \delta^+),$$

which implies that $A + \beta^- \delta^- V(0, \delta^-) < A + \beta^+ \delta^+ V(0, \delta^+)$. ■

Lemma 1 shows that with a typical tit-tat-for strategy (i.e., $q = 0$), the present bias makes it difficult to maintain mutual cooperation in a NE. Next, we investigate how the strategy's forgiveness (i.e., $q > 0$) affects the equilibrium. Because $V(q, \delta)$ is a continuous function in q at $q = 0$, the result in Lemma 1 implies that there exists $\overline{q} > 0$, such that for any $q \in [0, \overline{q})$,

$$\beta^- \delta^- V(q, \delta^-) < \beta^+ \delta^+ V(q, \delta^+). \tag{19}$$

For both (β^-, δ^-) and (β^+, δ^+), the values of discounted payoffs from mutual cooperations are the same as

$$C + \frac{\beta^- \delta^-}{1-\delta^-}C = C + \frac{\beta^+ \delta^+}{1-\delta^+}C. \tag{20}$$

From Eqs. (19-20), we know that the sustainability of mutual cooperation with (β^+, δ^+) is a sufficient condition for that with (β^-, δ^-) but not necessary.

B. Proof of Proposition 2

$\beta\delta V(q,\delta)$ can be written as $\frac{\beta\delta}{1-\delta}(1-\delta)V(q,\delta)$. Under equal patience, we obtain $\frac{\beta^-\delta^-}{1-\delta^-} = \frac{\beta^+\delta^+}{1-\delta^+}$. Therefore, inequality in (6) is identical to

$$\left(1-\delta^-\right)V(q^h,\delta^-) - \left(1-\delta^-\right)V(q^l,\delta^-) < \left(1-\delta^+\right)V(q^h,\delta^+) - \left(1-\delta^+\right)V(q^l,\delta^+) \qquad (21)$$

Defining function $f(q,\delta)$ as $(1-\delta)V(q,\delta)$, the inequality in Eq. (21) is identical to

$$f(q^h,\delta^-) - f(q^l,\delta^-) < f(q^h,\delta^+) - f(q^l,\delta^+), \qquad (22)$$

To establish the validity of the inequality in Eq. (22), we need to show that $\partial f(q,\delta)/\partial\delta > 0$ and $\partial^2 f(q,\delta)/(\partial q\partial\delta) > 0$ at $q = 0$. Differentiating $(1-\delta)V(q,\delta)$ with respect to δ and evaluating at $q = 0$, we obtain

$$\frac{\partial\,(1-\delta)V(q,\delta)}{\partial\delta} = \frac{A-Z}{(1+\delta)^2} > 0, \qquad (23)$$

Cross-differentiate $(1-\delta)V(q,\delta)$ with respect to δ and q and evaluate at $q = 0$, we have

$$\frac{\partial^2\,(1-\delta)V(q,\delta)}{\partial q\partial\delta} = \frac{C(1+\delta)^3 + Z(1-5\delta+\delta^2-\delta^3) + A(-2+2\delta-4\delta^2)}{(1-\delta)^2(1+\delta)^3} \qquad (24)$$

Because

$$2C > A + Z,$$

$$-(1+\delta)^3 = (1-5\delta+\delta^2-\delta^3) + (-2+2\delta-4\delta^2),$$

$$(1-5\delta+\delta^2-\delta^3) > (-2+2\delta-4\delta^2),$$

and

$$-2+2\delta-4\delta^2 < 0$$

we obtain $C(1+\delta)^3 + Z(1-5\delta+\delta^2-\delta^3) + A(-2+2\delta-4\delta^2) > 0$, which implies that

$$\frac{\partial^2\,(1-\delta)V(q,\delta)}{\partial q\partial\delta} > 0. \qquad (25)$$

From Eqs. (23) and (25), we know that as q increase, the higher value of δ (that is δ^+) makes $(1-\delta)V(q,\delta)$ increases faster than under the lower value of δ (i.e., δ^-). This means that with present bias, the increased amount of $\beta^-\delta^-V(q,\delta^-)$ from $q = q^l$ to any $q = q^h$ is greater than that of $\beta^+\delta^+V(q,\delta^+)$.

C. Proof of Proposition 3

Because $V(q, \delta)$ is a continuously differentiable function, the discrete result in Eq. (9) can directly support the continuous result in the Proposition.

D. Proof of Proposition 4

The proof is equivalent to show that

$$\beta^{+}\delta^{+}\left(D+\frac{\delta^{+}C}{1-\delta^{+}}\right) > \beta^{-}\delta^{-}\left(D+\frac{\delta^{-}C}{1-\delta^{-}}\right) \tag{26}$$

Since $\frac{\beta^{+}\delta^{+}}{1-\delta^{+}} = \frac{\beta^{-}\delta^{-}}{1-\delta^{-}}$, dividing $\frac{\beta^{+}\delta^{+}}{1-\delta^{+}}$ and $\frac{\beta^{-}\delta^{-}}{1-\delta^{-}}$ in both side of the inequality of Eq. (26), we have

$$\left(1-\delta^{+}\right) D + \delta^{+}C > \left(1-\delta^{-}\right) D + \delta^{-}C \tag{27}$$

Since $\delta^{+} > \delta^{-}$ and $D < C$, the inequalities in Eqs. (26-27) are true. This implies that the sustainability of mutual cooperation with (β^{+}, δ^{+}) is a sufficient condition for that with (β^{-}, δ^{-}) but not necessary.

E. Proof of Proposition 5

Under (β^{-}, δ^{-}), the condition for cooperation to be sustained in a NE under trigger strategy is

$$C + \frac{\beta^{-}\delta^{-}}{1-\delta^{-}}C \geq A + \frac{\beta^{-}\delta^{-}}{1-\delta^{-}}D \tag{28}$$

whereas under (β^{+}, δ^{+}), the condition is

$$C + \frac{\beta^{+}\delta^{+}}{1-\delta^{+}}C \geq A + \frac{\beta^{+}\delta^{+}}{1-\delta^{+}}D \tag{29}$$

Under equal patience, we have $\frac{\beta^{-}\delta^{-}}{1-\delta^{-}} = \frac{\beta^{+}\delta^{+}}{1-\delta^{+}}$. Thus, inequalities in Eqs. (28-29) are identical.

References

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution 24*(1), 3–25.

Axelrod, R. and W. D. Hamilton (1981). The evolution of cooperation. *Science 211*(4489), 1390–1396.

Barro, R. J. (1999). Ramsey meets laibson in the neoclassical growth model. *Quarterly Journal of Economics 114*(4), 1125–1152.

Benhabib, J., A. Bisin, and A. Schotter (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior 69*(2), 205–223.

Benzion, U., A. Rapoport, and J. Yagil (1989). Discount rates inferred from decisions: An experimental study. *Management science 35*(3), 270–284.

Cabo, F., G. Martin-Herran, and M. P. Martínez-García (2015). Non-constant discounting and Ak-type growth models. *Economics Letters 131*, 54–58.

Caliendo, F. N. and T. S. Findley (2014). Discount functions and self-control problems. *Economics Letters 122*(3), 416–419.

Chade, H., P. Prokopovych, and L. Smith (2008). Repeated games with present-biased preferences. *Journal of Economic Theory 139*(1), 157–175.

Dal Bó, P. and G. R. Fréchette (2019). Strategy choice in the infinitely repeated prisoner?s dilemma. *American Economic Review 109*(11), 3929–3952.

Gallo, E., Y. E. Riyanto, N. Roy, and T.-H. Teh (2022). Cooperation and punishment mechanisms in uncertain and dynamic social networks. *Games and Economic Behavior 134*, 75–103.

Imhof, L. A., D. Fudenberg, and M. A. Nowak (2007). Tit-for-tat or win-stay, lose-shift? *Journal of theoretical biology 247*(3), 574–580.

Kang, M. (2021). Aggregate savings under quasi-hyperbolic versus exponential discounting. *Economics Letters 207*, 110006.

Kang, M. (2025). Endogenous growth driven by government spending under quasi-hyperbolic discounting. *Economics Letters*, 112174.

Kang, M. and E. Kim (2024). Present bias and endogenous growth. *Economics Letters*, 111556.

Kim, J. (2023). The effects of time preferences on cooperation: Experimental evidence from infinitely repeated games. *American Economic Journal: Microeconomics 15*(1), 618–637.

Kirby, K. N. (1997). Bidding on the future: Evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General 126*(1), 54.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*(2), 443–477.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal 97*(387), 666–684.

Musau, A. (2014). Hyperbolic discount curves: a reply to ainslie. *Theory and Decision 76*, 9–30.

Myerson, J., L. Green, and M. Warusawitharana (2001). Area under the curve as a measure of discounting. *Journal of the Experimental Analysis of Behavior 76*(2), 235–243.

Nowak, M. and K. Sigmund (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature 364*(6432), 56–58.

Nowak, M. A. and K. Sigmund (1992). Tit for tat in heterogeneous populations. *Nature 355*(6357), 250–253.

Obara, I. and J. Park (2017). Repeated games with general discounting. *Journal of Economic Theory 172*, 348–375.

Schofield, N. (1985). Anarchy, altruism and cooperation: a review. *Social Choice and Welfare 2*(3), 207–219.

Strulik, H. (2015). Hyperbolic discounting and endogenous growth. *Economics Letters 126*, 131–134.

Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics letters 8*(3), 201–207.