Les Cahiers du GERAD ISSN: 0711–2440

An efficient payment scheme for sustaining cooperation in finitely repeated prisoner's dilemma games

E. Parilina, A. Pisareva, G. Zaccour

G–2025–62

September 2025

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2025
– Bibliothèque et Archives Canada, 2025

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2025
– Library and Archives Canada, 2025

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

An efficient payment scheme for sustaining cooperation in finitely repeated prisoner's dilemma games

Elena Parilina [a]
Alena Pisareva [a]
Georges Zaccour [b, c]

[a] *Saint Petersburg State University, Saint Petersburg, Russia*

[b] *Chair in Game Theory and Management, Montréal (Qc), Canada, H3T 1J4*

[c] *Department of Decision Sciences, HEC Montréal & GERAD, Montréal, (Qc), Canada, H3T 2A7*

e.parilina@spbu.ru
st055836@student.spbu.ru
georges.zaccour@gerad.ca

September 2025
Les Cahiers du GERAD
G–2025–62

Abstract : In this paper, we propose an efficient payment scheme for sustaining cooperation in finitely many times repeated Prisoner's Dilemma games. The scheme is part of a preplay arrangement that also defines the strategy that the two players should implement during the game. We consider two possibilities, namely, the limited retaliation strategy and the grim trigger strategy. Both strategies prescribe to play cooperatively, unless a defection is observed in which case both players switch to their noncooperative strategies, for an endogenously determined number of stages in the case of the limited retaliation strategy and until the end of the game in the case of the grim trigger strategy. We compare the results generated by the two strategies and determine the minimal number of stages needed to guarantee the existence of the payment scheme. We consider both discounted and undiscounted payoffs.

Keywords : Prisoner's dilemma; repeated games; minimal required savings; payment schemes; game duration

Résumé : Dans cet article, nous proposons un mécanisme de paiement efficace pour maintenir la coopération entre les parties dans un jeu de Dilemme du Prisonnier répétées un nombre fini de fois. Ce mécanisme fait partie d'un arrangement préalable qui définit également la stratégie que les deux joueurs doivent mettre en œuvre pendant la partie. Nous envisageons deux possibilités: la stratégie des représailles limitées et la stratégie du déclencheur sinistre. Les deux stratégies préconisent de coopérer, sauf en cas de défection, auquel cas les deux joueurs adoptent leurs stratégies non coopératives, pendant un nombre d'étapes déterminé de manière endogène pour la stratégie des représailles limitées et jusqu'à la fin de la partie pour la stratégie du déclencheur sinistre. Nous comparons les résultats engendrés par les deux stratégies et déterminons le nombre minimal d'étapes nécessaires pour garantir l'existence du système de paiement. Nous considérons les gains actualisés et non actualisés.

Mots clés : Dilemme du prisonnier; jeux répétés; épargne minimale requise; mécanismes de paiement; durée du jeu

1 Introduction

Suppose that a community of agents agree on a long-term contract aiming at optimizing a joint objective with a clause stipulating how to share the total gain. How can the players be sure that this cooperative agreement will remain in place until its maturity, i.e., each player will indeed implement the agreed-upon optimal decisions? The relevance of this question stems from the fact that, in general, the cooperative strategies are not part of a Nash equilibrium and therefore are not self-supported. A player will deviate from cooperation if she can get a better outcome-to-go by switching to a noncooperative strategy at any intermediate date during the game. This is nothing but individual rational behavior.

In the parlance of dynamic games, an agreement that breaks down before maturity is said to be time inconsistent. Of course, one can skip the problem by assuming that the contract is binding, but such thinking is wishful and lacks empirical support. Indeed, opting out of a cooperative arrangement after signing it is common in practice; the high rate of divorce is an illustration of that.

To start with, why players sign a long-term contract instead of keeping their options open by committing only one period at the time? First, negotiating to reach an acceptable arrangement is costly (in terms of dollars, time, and emotions); consequently, it is better to avoid repetitive negotiations whenever it is possible. Second, if the optimization problem that represents the situation at hand is truly dynamic, which happens anytime that history matters, then solving it as it were a series of static problems leads to suboptimal decisions. In short, long-term agreements are not a whim but a necessity in many instances.

The literature in dynamic games followed two general approaches to sustain cooperation (or collusion) over time. The first approach consists in embedding the cooperative solution with an equilibrium property. Such an approach has a long history in repeated games with the folk theorem, which (informally) states that if the players are sufficiently patient, then any Pareto-optimal outcome can be achieved as a Nash equilibrium using (typically) grim trigger strategies; see, e.g., Osborne and Rubinstein (1994) and Mailath and Samuelson (2006). Dutta (1995) established a similar result for stochastic games. Such strategies and approach have also been considered in differential and multistage games. For early contributions, see, e.g., Haurie and Tolwinski (1985), Tolwinski et al. (1986), Haurie and Pohjola (1987), and Haurie et al. (1994). An introduction to cooperative equilibria in differential games is available in Dockner et al. (2000) and Haurie et al. (2012).

The second approach aims at implementing time-consistent decomposition over time of the total cooperative payoff (allocation) such that, at any intermediate date, the cooperative payoff-to-go dominates (at least weakly) the noncooperative payoff-to-go for each player. This property is requested to hold along the cooperative state trajectory, which means that all players have been abiding by the agreement until the time of comparison. The early contributions in this area are Haurie (1976), Petrosyan (1977, 2008), and Petrosyan and Danilov (1979). Note that in this approach, either the agreement stands with all players in, or if falls apart, i.e., no agreement with a subset of players is considered. For a review of time-consistent solutions in cooperative dynamic games, we refer the interested reader to the surveys in Parilina et al. (2022); Petrosyan and Zaccour (2018); Yeung and Petrosyan (2012, 2018); Zaccour (2008), and the paper by Parilina and Tampieri (2018).

The folk theorems concern infinitely many times repeated games. Constructing a cooperative equilibrium in any class of finite-horizon dynamic games is generally out of reach because defecting in the last stage (or terminal time) is individually rational and this deviation cannot be punished. By a backward-induction argument, one can prove that the unique subgame-perfect equilibrium is then to play the Nash equilibrium controls at each time. Interestingly, this theoretical prediction has not always materialized in experiments where (at least some partial) cooperation is observed (see, e.g., Angelova et al. (2013)). The literature proposed various ideas to cope to maintain cooperation in finite-horizon dynamic games. See, e.g., Radner (1980), Benoit and Krishna (1985), Eswaran and Lewis (1986), Mailath et al. (2005), Flesch and Predtetchinski (2016), Flesch et al. (2022), and Parilina and Zaccour (2015).

In this paper, we take stock and extend the approach adopted in Parilina and Zaccour (2022, 2024), Pisareva and Parilina (2024), and in Parilina et al. (2025) to sustain cooperation in two-player finitely repeated Prisoner's Dilemma games. In Parilina and Zaccour (2022, 2024), the authors dealt with sustainability of cooperation in multistage games in deterministic and stochastic framework, respectively. In Pisareva and Parilina (2024), the behavior strategies with limited punishment resuming to cooperation after the end of punishment period are proposed. As here, Parilina et al. (2025) concerns cooperation in two-player Prisoner's Dilemma repeated games. The main idea in this publication is that the players agree on a preplay arrangement with two clauses. First, to implement a limited retaliation strategy (LRS) that dictates to play cooperatively, unless a defection is detected. If one or both players deviate from cooperation, then both players switch to their noncooperative strategy during a number of stages that is determined endogenously. Second, to adopt a payment scheme that allocates over stages the cooperative outcome. This scheme is based on a series of common-sense properties.

The contribution of the current paper with respect to the above cited ones is as follows:

1. In Parilina et al. (2025) the payment scheme differs most of the payment until the last stage to achieve efficiency. This is rather an extreme rewarding mechanism in a cooperative setting. Here, the gains of cooperation are allocated more generously during all stages.
2. We also consider the case where the players opt for the grim trigger strategy (GTS) instead of the LRS in the preplay arrangement. We compare the results in three respects:
   (a) The feasibility of a payment schedule.
   (b) The schedule of payments along the cooperative trajectory.
   (c) The minimum number of stages needed to sustain cooperation.
3. We state and analyze the results with discounted and undiscounted payoffs.

The rest of the paper is organized as follows: in Section 2, we introduce the model and limited retaliation strategies. The payment schemes and their properties are defined in Section 3, where we also determine minimally acceptable and efficient payment schemes. In Section 3.4, we examine how duration of the game influences feasibility conditions, i.e., the existence of the proposed payment schemes. Section 4 gives an illustrative example. We briefly conclude in Section 5.

2 Model

We consider a two-player Prisoner's Dilemma (PD) game repeated T times with the following stage game payoffs:

	C	N
C	(a, a)	(c, b)
N	(b, c)	(d, d)

where the actions C and N stand for cooperate and do not cooperate, respectively, and $b > a > d > c, 2a > c + b$. Denote by $t = 1, \dots, T$ the stage, and by $\rho \in (0, 1)$ the common discount factor satisfying the condition:[1]

$$\frac{b-a}{b-d} < \rho. \tag{1}$$

We assume that each player maximizes her total discounted payoff over the T stages. We shall also derive the results when the players do not discount their stream of payoffs. In this case, the condition in (1) is always satisfied.

[1] In infinitely many times repeated game, this condition guarantees that cooperation can be sustained forever with grim trigger strategies, i.e. the profile of such strategies admits a subgame perfect equilibrium with cooperative payoffs.

It is easy to check that the strategy repeating a pair of actions (N, N) in each stage is the unique subgame-perfect Nash equilibrium (for any discount factor ρ), while (C, C) in each stage is Pareto-optimal. Our aim is to design a preplay agreement that gives the players a higher payoff than the one obtained by playing (N, N) in each stage. Denote by $\mathcal{A} = \{C, N\}$ the set of actions in the stage game, by $A_t^i \in \mathcal{A}$ the action of player i at stage t, by $A_t = (A_t^1, A_t^2)$ profile of players' actions at stage t, and by $\pi_i(A_t)$ the payoff of player $i \in I = \{1, 2\}$ at stage t when the action profile A_t is played. A history $\mathcal{H}(t)$ of the game at stage t is the sequence of realized action profiles from all stages before t, that is,

$$\mathcal{H}(t) = (A_1, \dots, A_{t-1}), \quad \text{for } t = 2, 3, \dots,$$

with $\mathcal{H}(1) = \varnothing$.

The preplay agreement involves two items:

1. Play the history-dependent *limited retaliation strategy* (LRS or LR strategy) defined as follows (Pisareva and Parilina (2024), Parilina et al. (2025)):
   (a) Both players start by playing C.
   (b) Both players play C at t if $\mathcal{H}(t) = \big((C, C)_1, \dots, (C, C)_{t-1}\big)$, where $(C, C)_t$ is the action profile played in stage t.
   (c) If a first defection by either player is observed at t, that is, $A_t = (N, C)$, or $A_t = (C, N)$, or $A_t = (N, N)$, then both players switch to N at stage $t + 1$ for $m(t)$ stages.
      i. Resume cooperation at stage $m(t) + t + 2$, if $m(t) + t + 2 \le T$; otherwise, both players play N until the end of the game.
      ii. If $m(t) + t + 2 \le T$ and if for $\tau \ge m(t) + t + 2$, $A_\tau = (N, C)$, or $A_\tau = (C, N)$, or $A_\tau = (N, N)$, then both players play N starting from $\tau + 1$ until the end of the game.
2. Implement a *payment scheme*, which is efficient and stable against individual deviations, to allocate over stages the cooperative outcome.

We make two comments on this agreement. First, whereas the well-known grim trigger strategy (GTS or GT strategy) prescribes to play N until the end of the game after a first defection, the LRS forgives a first deviation, that is, cooperation can (possibly) resume after a noncooperative period. If a second defection occurs, then LRS and GTS coincide in prescribing to play N until the end. We shall illustrate the differences between the two strategies later on. Second, we shall state in the next section some desirable properties that the payment scheme should satisfy. Importantly, the rewards defined by the payment scheme are not equilibrium outcomes; they are designed to incentivize the players to cooperate throughout the game, i.e., they prevent players' deviations along the cooperative trajectory.

To determine the retaliation period $m(t)$, we follow the same steps as in Parilina et al. (2025). Suppose that Player i deviates from cooperation at stage t and Player j does not.

If Player j plays C after the end of the retaliation period until the end of the game, then she gets

$$X_j = \sum_{\tau = t + m(t) + 1}^{T} \rho^{\tau - 1} a.$$

If she plays N, then she obtains

$$Y_j = \rho^{t + m(t)} b + \sum_{\tau = t + m(t) + 2}^{T} \rho^{\tau - 1} d.$$

Now, if Player i plays N during the stages t to $t + m(t)$, then her gain is given by

$$W_i = \rho^{t-1} b + \sum_{\tau = t+1}^{t + m(t)} \rho^{\tau - 1} d,$$

and by

$$V_i = \sum_{\tau=t}^{t+m(t)} \rho^{\tau-1} a,$$

if she plays C.

To determine the retaliation period $m(t)$, we adopt the following two rules:

R1: $X_j \geq Y_j$. Player j should be better off resuming cooperation after the retaliation period.

R2: $V_i \geq W_i$. During the retaliation period, Player i's gain should be at most equal to what she could get if (C, C) were played.

Let

$$Z = \log_\rho \left(\frac{a - b + \rho(b-d)}{a-d} \right).$$

After some straightforward manipulations (see Parilina et al. (2025) for the details), rules R1 and R2 lead to

$$Z - 1 \leq m(t) \leq T - t - Z. \tag{2}$$

Two cases must be considered:

1. If $Z - 1 \geq T - t - Z$, then the duration of the retaliation period $m(t)$ exceeds the number of remaining stages, and consequently the game will proceed noncooperatively from $t+1$ until the end.
2. If $Z - 1 \leq T - t - Z$, which is equivalent to

$$T - t + 1 \geq 2Z, \tag{3}$$

   then $m(t)$ can take any value between the two bounds in (2). Observe that the upper bound depends on both the stage at which the defection happens and on the terminal stage, while the lower bound is a constant defined by the payoffs in the stage game and the discount factor.

For the duration of the retaliation period to be an integer number of stages, we let

$$\underline{m} = \lceil Z - 1 \rceil, \tag{4}$$

$$\overline{m} = \lfloor T - 1 - Z \rfloor, \tag{5}$$

where $\lceil Z - 1 \rceil$ denotes the rounding up to the next integer of $Z - 1$, and $\lfloor T - 1 - Z \rfloor$ denotes the rounding down to the previous integer. Observe that $\overline{m} = m(1)$, i.e., the duration of retaliation if defection happens at stage $t = 1$. Further, from (2), we have that the maximal duration of retaliation for any $t = 2, \ldots, T$ satisfies the following recurrence:

$$\overline{m}(t) = \overline{m} - t + 1.$$

If the players do not discount their streams of payoffs, then the sum of gains defined above become

$$\tilde{X}_j = a\left(T - (t + \tilde{m}(t) + 1)\right), \qquad \tilde{Y}_j = b + d\left(T - (t + \tilde{m}(t) + 2)\right),$$

$$\tilde{V}_i = a \cdot \tilde{m}(t), \qquad \tilde{W}_i = b + \tilde{m}(t)\, d,$$

where we tilted the variables to distinguish with the notation in the discounted case. It can be easily checked that the duration of retaliation is bounded as follows:

$$\frac{b-a}{a-d} \leq \tilde{m}(t) \leq T - t - \frac{b-a}{a-d} - 1. \tag{6}$$

Again, we have two cases:

1. If $\frac{b-a}{a-d} \geq T - t - \frac{b-a}{a-d} - 1$, then $\tilde{m}(t)$ exceeds the number of remaining stages, and the game will be played noncooperatively from $t+1$ until the end.
2. If $\frac{b-a}{a-d} \leq T - t - \frac{b-a}{a-d} - 1$, then $\tilde{m}(t)$ is defined by (6).

3 Payment schemes

As the action profile (C, C) in each stage is not part of a subgame-perfect Nash equilibrium under LRS (it is easy to verify this), our aim is to design a payment scheme (PS) to incentivize the two players to cooperate throughout the whole game. Intuitively, any such PS should satisfy some accounting principles and be individual rational, i.e., each player prefers the payment scheme to its absence.

Definition 1. A payment scheme $\mathcal{P}$ is a set of payments $(p_i(t) : i \in I, t \in \mathbf{T})$ that player i receives in stage $t \in \mathbf{T} = \{1, \ldots, T\}$ along the cooperative trajectory (choosing action profile $(C; C)$ at each stage of the game) when both players adopt the LR strategy.

Denote by $s\,(t)$ the savings made at stage t for future use. It is defined by

$$s\,(t) = \sum_{i\in I} (\pi_i(C, C) - p_i(t)) = 2a - p_1(t) - p_2(t)$$

be the savings at stage t. Capitalizing the stream of savings at the same discount factor ρ leads to the following cumulative savings $S(t)$ at stage $t = 2, \ldots, T$:

$$S(t) = \sum_{\tau=1}^{t-1} \frac{1}{\rho^{t-\tau}} s\,(\tau) = \sum_{\tau=1}^{t-1} \frac{1}{\rho^{t-\tau}} \left(2a - p_1(\tau) - p_2(\tau)\right),$$

and $S(1) = 0$.

Parilina and Zaccour (2022, 2024) and Parilina et al. (2025) considered the following feasibility, efficiency, and stability properties, and the minimal required savings property in Parilina and Zaccour (2022):

Feasibility: $\mathcal{P}$ is feasible if the sum of payments is less or equal to the sum of gains plus available savings, i.e.,

$$\sum_{i\in I} p_i(t) \le \sum_{i\in I} \pi_i(C, C) + S(t) = 2a + S(t), \quad \text{for } t = 1, \ldots, T. \tag{7}$$

Minimal required savings: $\mathcal{P}$ satisfies the minimal required savings (MRS) property if

$$\begin{cases} S(T+1) & = 0, \\ S(t) & = \left[(2a - \sum_{i=1}^{2} p_i(t)) - \rho S(t+1)\right]^+, \text{ for } t = 1, \ldots, T, \end{cases} \tag{8}$$

where $[A]^+ = A$ if $A > 0$, and $[A]^+ = 0$ if $A \le 0$.

Efficiency: $\mathcal{P}$ is efficient if the total discounted payments to the players is equal to the total discounted payoffs when the players choose (C, C) in each stage. Formally,

$$\sum_{i\in I} \sum_{t=1}^{T} \rho^{t-1} p_i(t) = \sum_{i\in I} \sum_{t=1}^{T} \rho^{t-1} \pi_i(C, C) = 2a \frac{1-\rho^T}{1-\rho}. \tag{9}$$

Stability: $\mathcal{P}$ is stable against individual deviation (SAID) if

$$\sum_{\tau=t}^{T} \rho^{\tau-t} p_i(\tau) \ge BR_i(t), \; i = 1, 2 \text{ and } 1 \le t \le T, \tag{10}$$

where $BR_i(t)$ is the *best reward* that Player i realizes if she deviates from C to N at stage t, anticipating that both players implement LR strategies.

It is easy to verify that for LR strategy, we have

$$BR_i(t) = \begin{cases} b + \rho \begin{cases} \frac{d(1-\rho^{m(t)})}{1-\rho} + \frac{a\rho^{m(t)}(1-\rho^{T-m(t)-t})}{1-\rho}, & T-t+1 \geq 2Z, \\ \frac{d(1-\rho^{T-t})}{1-\rho}, & T-t+1 < 2Z, \end{cases} \\ \text{for } 1 \leq t < T, \\ \\ b, \quad \text{for } t = T, \end{cases} \tag{11}$$

where $m(t)$ satisfies condition (2).

Feasibility is an accounting property that must be satisfied in the absence of subsidies. The MRS property adds some farsightedness in the process by having some provisions for future use. Efficiency is desirable in the sense that it does not leave money on the table. SAID is a basic requirement from an individual rationality point of view. No player would accept an allocation that gives her less than what it can secure by acting selfishly, i.e., switching from a cooperative trajectory. While it is not granted, this property is intuitively easier to satisfy when all conditions in (10) are equalities, which leads to the idea of a minimally acceptable payment scheme (MAPS).

Definition 2. A minimally acceptable payment scheme $\mathcal{P}^{ma}$ is given by

$$\mathcal{P}^{ma} = \left\{ p_i^{ma}(t) : \sum_{\tau=t}^{T} \rho^{\tau-t} p_i^{ma}(\tau) = BR_i(t),\ i = 1, 2 \text{ and } 1 \leq t \leq T \right\}, \tag{12}$$

for which the feasibility property is satisfied.

In the undiscounted case, the cumulative savings are given by[2]

$$\tilde{S}(t) = \sum_{\tau=1}^{t-1} \left(2a - \tilde{p}_1(\tau) - \tilde{p}_2(\tau)\right).$$

The efficiency property becomes

$$\sum_{i\in I} \sum_{t=1}^{T} \tilde{p}_i(t) = \sum_{i\in I} \sum_{t=1}^{T} \pi_i(C, C) = 2aT, \tag{13}$$

and the best reward under LR strategy reads as follows:

$$\widetilde{BR}_i(t) = \begin{cases} b + \begin{cases} \tilde{m}(t)d + a(T - \tilde{m}(t) - t), & T-t-1 \geq 2\frac{b-a}{a-d}, \\ (T-t)d, & T-t-1 < 2\frac{b-a}{a-d}, \end{cases} \\ \text{for } 1 \leq t < T, \\ \\ b, \quad \text{for } t = T, \end{cases} \tag{14}$$

where $\tilde{m}(t)$ satisfies condition (6).

In the undiscounted case, the MAPS definition becomes

$$\tilde{\mathcal{P}}^{ma} = \left\{ \tilde{p}_i^{ma}(t) : \sum_{\tau=t}^{T} \tilde{p}_i^{ma}(\tau) = \widetilde{BR}_i(t),\ i = 1, 2 \text{ and } 1 \leq t \leq T \right\}, \tag{15}$$

and the feasibility condition given by inequality: $\sum_{i\in I} \tilde{p}_i(t) \leq 2a + \tilde{S}(t)$ for any $t = 1, \ldots, T$ should be satisfied for $\tilde{\mathcal{P}}^{ma}$.

[2] We use tilde to distinguish variables in undiscounted case in comparison with discounted case, e.g. payment to player i in stage t in discounted and undiscounted cases are denoted by $p_i(t)$ and $\tilde{p}_i(t)$, respectively.

Remark 1. If both players use the grim trigger strategy instead of the LRS, then the best reward defined in (11) becomes:

1. In the discounted case,

$$BR_i^G(t) = \begin{cases} b + d\rho\frac{1-\rho^{T-t}}{1-\rho}, & \text{for } t = 1, \dots, T-1, \\ b, & \text{for } t = T, \end{cases}$$

   which coincides with $BR_i(t)$ when $T - t + 1 < 2Z$, that is, when the duration of the retaliation period exceeds the remaining number of stages in the game.

2. In the undiscounted case,

$$\widetilde{BR}_i^G(t) = \begin{cases} b + d(T-t), & \text{for } t = 1, \dots, T-1, \\ b, & \text{for } t = T, \end{cases}$$

   which corresponds to $\widetilde{BR}_i(t)$ when $T - t - 1 < 2\frac{b-a}{a-d}$, that is, when the duration of the retaliation period exceeds the remaining number of stages in the game.

We make three comments. First, the existence of a MAPS is not granted. One aim of this section is to establish the conditions under which it exists. Second, as the best reward that a player can realize at any t depends on the strategy, it is clear that the payments will not be (necessarily) the same under the LR and GT strategies. We shall determine and contrast the payments induced by these two strategies. Third, the best reward depends on the duration of the retaliation period. Although any period lying between the lower and upper bounds established in (2) for the discounted case and in (6) for the undiscounted case is admissible, we shall focus our analysis on the upper bound, which depends on the defection stage under the LR strategy.

3.1 Minimally acceptable payment scheme with LR strategy

3.1.1 Discounted payoffs

This case has been considered, to a large extent but not fully, in Parilina et al. (2025), so here we only provide the formulae obtained in that paper.

Let $\lambda - 1$ be the last stage at which the set of retaliation duration is non-empty. The value is the unique solution of the equation $\underline{m} = \overline{m}(\lambda - 1)$.

1. A MAPS with LR strategy with maximal retaliation duration, i.e., $\mathcal{P}^{ma} = (p_i^{ma}(t) : i \in I, t \in \mathbf{T})$, is as follows:

$$p_i^{ma}(t) = p^{ma}(t) = \begin{cases} b, & \text{if } t = T; \\ b(1-\rho) + \rho d, & \text{if } T + 1 - 2Z < t < T; \\ b(1-\rho) + \rho d + \frac{(a-d)(\rho^{\overline{m}-t+2} - \rho^{T-t+1})}{1-\rho}, & \text{if } T - 2Z < t \le T - 2Z + 1; \\ b(1-\rho) + \rho d, & \text{if } 1 \le t \le T - 2Z. \end{cases} \tag{16}$$

2. $\mathcal{P}^{ma}$ exists if the following condition holds:

$$\rho - \frac{(a-d)}{(b-d)} \frac{\rho^{\overline{m}+1} - \rho^T}{1 - \rho^{\lambda-1}} \ge \frac{b-a}{b-d}. \tag{17}$$

Recalling the condition $\frac{b-a}{b-d} < \rho$ imposed from the outset on the game data, it is clear that the above condition is more restrictive, which makes nontrivial the existence of the payment scheme $\mathcal{P}^{ma}$.

3.1.2 Undiscounted payoffs

To construct a minimally acceptable payment scheme with punishment duration $\tilde{m}(t)$, we proceed backward starting with period $t = T$.

1. At $t = T$, set $\tilde{p}_i(T) = \widetilde{BR}_i(T) = b \ \forall i \in I$.

2. At $t = T-1$, $\tilde{p}_i(T-1) = \widetilde{BR}_i(T-1) - \widetilde{BR}_i(T) = b + d - b = d$ for any $i \in I$, since the punishment cannot be realized at $T-1$ because $0 = T-(T-1)-1 < 2\frac{b-a}{a-d}$.

3. Let $t = T-2, \ldots, 1$, then
$\tilde{p}_i(t) = \widetilde{BR}_i(t) - \sum_{\tau=t+1}^{T} \tilde{p}_i(\tau) = \widetilde{BR}_i(t) - \widetilde{BR}_i(t+1)$.
Consider in turn the three cases:

$t \in \{\lambda, \ldots T-2\}$, where $T - t - 1 < 2\frac{b-a}{a-d}$, i.e., there is no punishment satisfying (2):

$$\tilde{p}_i(t) = \widetilde{BR}_i(t) - \widetilde{BR}_i(t+1) = b + (T-t)d - (b + (T-t-1)d) = d.$$

$t = \lambda - 1$:

$$\begin{aligned}\tilde{p}_i(t) = \widetilde{BR}_i(\lambda-1) - \widetilde{BR}_i(\lambda) = \\ &= b + \tilde{m}(t)d + a(T - \tilde{m}(t) - t) - (b + (T-t-1)d) = \\ &= \tilde{m}(t)d + a(T - \tilde{m}(t) - t) - (T-t-1)d \\ &= (T - \tilde{m}(t) - t)(a-d) + d.\end{aligned}$$

$t \in \{1, \ldots, \lambda-2\}$, where $T - t - 1 \geq 2\frac{b-a}{a-d}$, i.e., there is a punishment satisfying (2):

$$\begin{aligned}\tilde{p}_i(t) = \widetilde{BR}_i(t) - \widetilde{BR}_i(t+1) = \\ &= b + \tilde{m}(t)d + a(T - \tilde{m}(t) - t) - \Big(b + \tilde{m}(t+1)d + a(T - \tilde{m}(t+1) - t - 1)\Big) \\ &= a + (d-a)\Big(\tilde{m}(t) - \tilde{m}(t+1)\Big).\end{aligned}$$

Substituting the maximal retaliation duration in LR strategy, we obtain

$$\tilde{\mathcal{P}}^{ma} = (\tilde{p}_i^{ma}(t) : i \in I, t \in \mathbf{T}),$$

where

$$\tilde{p}_i^{ma}(t) = \tilde{p}^{ma}(t) = \begin{cases} b, & \text{if } t = T, \\ d, & \text{if } T - 1 - 2\frac{b-a}{a-d} < t < T, \\ (T - \tilde{\bar{m}} - 1)(a-d) + d, & \text{if } T - 2\frac{b-a}{a-d} - 2 < t \leq T - 2\frac{b-a}{a-d} - 1, \\ d, & \text{if } 1 \leq t \leq T - 2\frac{b-a}{a-d} - 2, \end{cases} \tag{18}$$

where $\tilde{\bar{m}} = \lfloor T - \frac{b-a}{a-d} - 2 \rfloor$.

Proposition 1. In the undiscounted case, $\tilde{\mathcal{P}}^{ma}$ exists if the following condition is satisfied:

$$\left\lfloor T - 2 - \frac{b-a}{a-d} \right\rfloor \geq 2\frac{b-a}{a-d} + 1.$$

Proof. Consider the following cases:

1. For stage $t \leq T - 2\frac{b-a}{a-d} - 2$, the feasibility condition is satisfied because $d < a$.

2. For the stage t such that $T - 2\frac{b-a}{a-d} - 2 < t \leq T - 2\frac{b-a}{a-d} - 1$, we have:

$$\tilde{S}(t) = 2\sum_{\tau=1}^{t-1}(a-d) = 2(a-d)(t-1)$$

and the feasibility condition, that is,

$$\tilde{p}_1^{ma}(t) + \tilde{p}_2^{ma}(t) \leq 2a + \tilde{S}(t) = 2a + 2(a-d)(t-1),$$

is met since

$$\begin{aligned} 2\Big((T - \tilde{\overline{m}} - 1)(a-d) + d\Big) &\leq 2a + 2(a-d)(t-1), \\ (T - \tilde{\overline{m}} - 1)(a-d) &\leq (a-d)t, \\ T - \tilde{\overline{m}} - 1 - t &\leq 0 \text{ if} \\ T - \left\lfloor T - \frac{b-a}{a-d} - 2 \right\rfloor - 1 - (T - 2\frac{b-a}{a-d} - 2) &\leq 0. \end{aligned}$$

3. For stage t such that $T - 1 - 2\frac{b-a}{a-d} < t < T$:

$$\begin{aligned} 2d &\leq 2a + 2(a-d)(t-2) + 2a - 2\Big((T - \tilde{\overline{m}} - 1)(a-d) + d\Big), \\ 2(a-d)(t-1) + 2a &\geq 2\Big((T - \tilde{\overline{m}} - 1)(a-d) + d\Big), \end{aligned}$$

and we come the previous case.

4. For $t = T$, the feasibility condition is satisfied if:

$$\begin{aligned} 2b &\leq 2a + 2(a-d)(T-2) + 2a - 2\Big((T - \tilde{\overline{m}} - 1)(a-d) + d\Big), \\ (b-a) &\leq (a-d)(T-1) - \Big((T - \tilde{\overline{m}} - 1)(a-d)\Big), \\ (b-a) &\leq (a-d)(T-1) - (a-d)(T-1) + \tilde{\overline{m}}(a-d), \\ \frac{b-a}{a-d} &\leq \tilde{\overline{m}}, \end{aligned}$$

which is fulfilled by R2.

This finishes the proof. □

3.2 Minimally acceptable payment scheme with GT strategy

Denote by $\mathcal{P}^{ma,G} = (p_i^{ma,G}(t) : i \in I, t \in \mathbf{T})$ a minimally acceptable payment scheme when the players use the GT strategy. In the discounted case, Parilina and Zaccour (2022) determined this MAPS as follows:

$$p_i^{ma,G}(t) = p^{ma,G}(t) = \begin{cases} b, & \text{if } t = T, \\ (1-\rho)b + \rho d, & \text{if } T-1 \leq t \leq 1, \end{cases} \tag{19}$$

with the following feasibility condition:

$$\frac{\rho - \rho^T}{1 - \rho^T} \geq \frac{b-a}{b-d}. \tag{20}$$

To define the payment scheme in the undiscounted case, it suffices to set $\rho = 1$ in (19) to obtain

$$\tilde{p}_i^{ma,G}(t) = \tilde{p}^{ma,G}(t) = \begin{cases} b, & \text{if } t = T, \\ d, & \text{if } T-1 \leq t \leq 1, \end{cases} \tag{21}$$

and the following feasibility condition:

$$T \geq \frac{b-d}{a-d}. \tag{22}$$

3.3 Efficient payment scheme

A MAPS is never efficient. To prove this statement, it suffices to sum the payments over the T stages and notice that the result does not satisfy the efficiency property. One intuitive way to enforce cooperation throughout the game is to hold (most of) the payment until T. This is the approach followed in Parilina et al. (2025). However, such trivial scheme may not be desirable as the players may want to receive higher payoffs during the game.

Let $\mathcal{P}^{mrs} = (p_i^{mrs}(t) : i \in I, t \in \mathbf{T})$ be an efficient payment scheme that satisfies the MRS property (see (8)) in the discounted case.

Proposition 2. If a minimally acceptable payment scheme $\mathcal{P}^{ma}$ exists, then the scheme

$$p_i^{mrs}(t) = \begin{cases} p_i^{ma}(t), & \text{if } t = T, \\ p_i^{ma}(t), & \text{if } \varepsilon_i(t) < 0 \text{ and } 1 \leq t < T, \\ p_i^{ma}(t) + \varepsilon_i(t), & \text{if } \varepsilon_i(t) \geq 0 \text{ and } 1 \leq t < T, \end{cases} \tag{23}$$

where

$$\varepsilon_i(t) = \frac{\sum\limits_{\tau=t+1}^{T} (a - p_i^{mrs}(\tau))\rho^{\tau-1} + \rho^{t-1}(a - p_i^{ma}(t))}{\rho^{t-1}},$$

exists and satisfies the Feasibility, SAID, Efficiency, and the MRS properties.

Proof. Suppose that $\mathcal{P}^{ma}$ exists and denote by $S^{mrs}(t)$ the savings for $\mathcal{P}^{mrs}$, and by $S(t)$ the savings for $\mathcal{P}^{ma}$. We construct a payment scheme $\mathcal{P}^{mrs}$ backwards from $t = T$ to $t = 1$ as follows:

$t = T$**:** At the terminal stage, $p_i^{mrs}(T) = p_i^{ma}(T) = BR_i(T) = b$, and the payoff of any player on the cooperative trajectory is equal to a. To satisfy feasibility, the necessary and sufficient savings for the last stage should be $S^{mrs}(T) = 2(b-a)$ (it is obtained taking into account that $S^{mrs}(T+1) = 0$ since the efficiency property should be satisfied, and $S^{mrs}(t) = \rho S^{mrs}(t+1) - (2a - p_1^{mrs}(t) - p_2^{mrs}(t)))$. By construction $S^{mrs}(T+1) = 0$, then the feasibility at T is satisfied.

$t = T-1, \ldots, 1$**:** Let us define the payments along the cooperative trajectory by equation (23), that is,

$$p_i^{mrs}(t) = \begin{cases} p_i^{ma}(t), & \text{if } \varepsilon_i(t) < 0, \\ p_i^{ma}(t) + \varepsilon_i(t), & \text{if } \varepsilon_i(t) \geq 0, \end{cases}$$

where $\varepsilon_i(t) = \frac{\sum\limits_{\tau=t+1}^{T} (a - p_i^{mrs}(\tau))\rho^{\tau-1} + \rho^{t-1}(a - p_i^{ma}(t))}{\rho^{t-1}}$.

The value $\varepsilon_i(t)$ consists of two terms. The first one is the saving and the second is the difference between the (cooperative) payoff in the current stage t and the payment.

By construction the MRS property is satisfied. We need to verify the feasibility defined by (23) at stages: $t = T-1, \ldots, 1$. Notice that the feasibility condition at stage t is equivalent to $S^{mrs}(t+1) \geq 0$. First, we rewrite the formula defining $\varepsilon_i(t)$ (when players are symmetric):

$$\varepsilon_i(t) = \frac{\sum_{\tau=t+1}^{T} (a - p_i^{mrs}(\tau))\rho^{\tau-1} + \rho^{t-1}(a - p_i^{ma}(t))}{\rho^{t-1}}$$

$$
\begin{aligned}
&= \sum_{\tau=t+1}^{T} (a - p_i^{mrs}(\tau))\rho^{\tau-t} + (a - p_i^{ma}(t)) \\
&= (a - p_i^{ma}(t)) + \frac{1}{\rho^t}\Big(\sum_{\tau=1}^{T}(a - p_i^{mrs}(\tau))\rho^{\tau} - \sum_{\tau=1}^{t}(a - p_i^{mrs}(\tau))\rho^{\tau}\Big) \\
&= (a - p_i^{ma}(t)) + \frac{\Big(S^{mrs}(T+1) - S^{mrs}(t+1)\Big)}{2\rho^t} \\
&= (a - p_i^{ma}(t)) - \rho S^{mrs}(t+1)/2.
\end{aligned}
$$

Second, starting from $T-1$, we verify backward the feasibility step by step:

1. $t = T-1$: since $S^{mrs}(T) = 2(b-a) > 0$, feasibility at $T-1$ holds true.
   ...
2. $t-1$. Let feasibility be satisfied for t. Prove it for stage $t-1$. To do this, we consider the case when

$$\varepsilon_i(t) < 0,$$

and rewrite this condition as follows:

$$
\begin{aligned}
(a - p_i^{ma}(t)) - \rho S^{mrs}(t+1)/2 < 0,& \\
(a - p_i^{ma}(t)) < \rho S^{mrs}(t+1)/2,& \\
(a - p_i^{ma}(t)) < S^{mrs}(t)/2 + a - p_i^{mrs}(t),& \\
0 < S^{mrs}(t)/2,&
\end{aligned}
$$

since $p_i^{mrs}(t) = p_i^{ma}(t)$, if $\varepsilon_i(t) < 0$. Therefore, feasibility at $t-1$ holds true.
Now consider the case when

$$\varepsilon_i(t) \geq 0,$$

from which it follows:

$$
\begin{aligned}
p_i^{mrs}(t) &= p_i^{ma}(t) + \varepsilon_i(t), \\
p_i^{mrs}(t) &= p_i^{ma}(t) + \frac{\sum_{\tau=t+1}^{T}(a - p_i^{mrs}(\tau))\rho^{\tau-1} + \rho^{t-1}(a - p_i^{ma}(t))}{\rho^{t-1}}, \\
0 &= \sum_{\tau=t+1}^{T} (a - p_i^{mrs}(\tau))\rho^{\tau-t} + (a - p_i^{mrs}(t)), \\
0 &= (a - p_i^{mrs}(t)) - \rho S^{mrs}(t+1)/2, \\
\rho S^{mrs}(t+1)/2 &= (a - p_i^{mrs}(t)), \\
(a - p_i^{mrs}(t)) + S^{mrs}(t)/2 &= (a - p_i^{mrs}(t)), \\
S^{mrs}(t)/2 &= 0,
\end{aligned}
$$

which proves that feasibility at $t-1$ holds.

The payment scheme $\mathcal{P}^{mrs} = \{p_i^{mrs}(t), i \in \{1,2\}, 1 \leq t \leq T\}$ defined by (23) is stable against individual deviation because $p_i^{mrs}(t) \geq p_i^{ma}(t)$ and SAID property is satisfied for minimally acceptable payment scheme $\{p_i^{ma}(t)\}$. □

Let $\tilde{\mathcal{P}}^{mrs} = (\tilde{p}_i^{mrs}(t) : i \in I, t \in \mathbf{T})$ be an efficient payment scheme that satisfies the MRS property in the undiscounted case.

Proposition 3. If the minimally acceptable payment scheme $\tilde{\mathcal{P}}^{ma}$ exists, then the scheme

$$\tilde{p}_i^{mrs}(t) = \begin{cases} \tilde{p}_i^{ma}(t), & \text{if } t = T, \\ \tilde{p}_i^{ma}(t), & \text{if } \tilde{\varepsilon}_i(t) < 0 \text{ and } 1 \le t < T, \\ \tilde{p}_i^{ma}(t) + \tilde{\varepsilon}_i(t), & \text{if } \tilde{\varepsilon}_i(t) \ge 0 \text{ and } 1 \le t < T, \end{cases} \tag{24}$$

where

$$\tilde{\varepsilon}_i(t) = \sum_{\tau=t+1}^{T} (a - \tilde{p}_i^{mrs}(t)) + (a - \tilde{p}_i^{ma}(t)),$$

exists and satisfies the Feasibility, SAID, Efficiency, and the MRS properties.

Proof. The proof is similar to the proof of Proposition 2 with corrections made for the absence of discounting in players' payoffs. □

The two propositions provide the formulae for the payments that the players need to use to achieve an efficient outcome in the finitely many times repeated Prisoner's Dilemma game.

Remark 2. The above results hold under both the limited retaliation and the grim trigger strategies.

3.4 Impact of game duration on payment scheme existence

To wrap up, designing a PS, supported by either LR or GT strategies, which incentivizes the players to choose (C, C) in all stages and realizes the efficient outcome boils down to the existence of a MAPS. As we have seen, the conditions for this existence depends on all parameters involved in the game, i.e., the numbers in the stage-payoff matrix, the duration of the game, and on the discount factor.

In this subsection, we look at the relationships between the duration of the game and the existence of a MAPS. Also, we compare the results under the LR and GT strategies.

3.4.1 Discounted case

The feasibility conditions under GTS and LRS are, respectively, as follows:

$$\frac{\rho - \rho^T}{1 - \rho^T} \ge \frac{b - a}{b - d}, \tag{25}$$

$$\rho - \frac{(a - d)}{(b - d)} \frac{\rho^{\overline{m}+1} - \rho^T}{1 - \rho^{\lambda-1}} \ge \frac{b - a}{b - d}, \tag{26}$$

where $\overline{m} = \lfloor T - 1 - Z \rfloor$. Let

$$f_{GTS}(T) = \frac{\rho - \rho^T}{1 - \rho^T}, \tag{27}$$

$$f_{LRS}(T) = \rho - \frac{(a - d)}{(b - d)} \frac{\rho^{\overline{m}+1} - \rho^T}{1 - \rho^{\lambda-1}}. \tag{28}$$

Proposition 4. The following holds true:

1. Both $f_{LRS}(T)$ and $f_{GTS}(T)$ are increasing in T.
2. For all parameter values,
$$\lim_{T \to \infty} f_{GTS}(T) = \lim_{T \to \infty} f_{LRS}(T) = \rho.$$
3. For all parameter values,
$$f_{GTS}(T) \ge f_{LRS}(T).$$

Proof.

1. Consider the difference

$$f_{GTS}(t+1) - f_{GTS}(t) = \frac{\rho - \rho^{t+1}}{1-\rho^{t+1}} - \frac{\rho - \rho^t}{1-\rho^t} = \frac{\rho^t - 2\rho^{t+1} + \rho^{t+2}}{(1-\rho^{t+1})(1-\rho^t)}$$
$$= \frac{\rho^t(1-2\rho+\rho^2)}{(1-\rho^{t+1})(1-\rho^t)} = \frac{\rho^t(1-\rho)^2}{(1-\rho^{t+1})(1-\rho^t)},$$

which is clearly positive for any t.

Now, consider the difference[3]

$$f_{LRS}(t+1) - f_{LRS}(t) = \rho - \frac{a-d}{b-d}\frac{\rho^{\overline{m}(t+1)+1} - \rho^{t+1}}{1-\rho^{\lambda(t+1)-1}} - \rho + \frac{a-d}{b-d}\frac{\rho^{\overline{m}(t)+1} - \rho^t}{1-\rho^{\lambda(t)-1}}.$$

Showing that $f_{LRS}(t+1) - f_{LRS}(t) > 0$ is equivalent to have:

$$\frac{a-d}{b-d}\frac{\rho^{\overline{m}(t)+2} - \rho^{t+1}}{1-\rho^{\lambda(t+1)-1}} - \frac{a-d}{b-d}\frac{\rho^{\overline{m}(t)+1} - \rho^t}{1-\rho^{\lambda(t)-1}} < 0 \Leftrightarrow$$
$$\frac{\rho^{\overline{m}(t)+2} - \rho^{t+1}}{1-\rho^{\lambda(t+1)-1}} - \frac{\rho^{\overline{m}(t)+1} - \rho^t}{1-\rho^{\lambda(t)-1}} < 0 \Leftrightarrow$$
$$(\rho^{\overline{m}(t)+2} - \rho^{t+1})(1-\rho^{\lambda(t)-1}) - (\rho^{\overline{m}(t)+1} - \rho^t)(1-\rho^{\lambda(t+1)-1}) < 0 \Leftrightarrow$$
$$(\rho-1)(\rho^{\overline{m}(t)+1} - \rho^t) < 0 \Leftrightarrow$$
$$\overline{m}(t) + 1 < t,$$

which holds true for any t.

2. As $\rho \in (0,1)$, it is straightforward to get

$$\lim_{t\to\infty} f_{GTS}(t) = \lim_{t\to\infty} \frac{\rho-\rho^t}{1-\rho^t} = \rho,$$
$$\lim_{t\to\infty} f_{LRS}(t) = \lim_{t\to\infty}\left(\rho - \frac{(a-d)}{(b-d)}\frac{\rho^{\overline{m}(t)+1} - \rho^T}{1-\rho^{\lambda(t)-1}}\right) = \rho.$$

3. Consider the difference

$$f_{GTS}(T) - f_{LRS}(T) = \frac{\rho-\rho^T}{1-\rho^T} - \left(\rho - \frac{(a-d)}{(b-d)}\frac{\rho^{\overline{m}+1} - \rho^T}{1-\rho^{\lambda-1}}\right).$$

Recall that

$$Z = \log_\rho\left(\frac{a-b+\rho(b-d)}{a-d}\right),$$

so $\rho^Z = \frac{a-b+\rho(b-d)}{a-d}$ and $\frac{a-d}{b-d} = \frac{1-\rho}{1-\rho^Z}$. Substitute for $\frac{a-d}{b-d} = \frac{1-\rho}{1-\rho^Z}$ into $f_{GTS}(T) - f_{LRS}(T)$ to get

$$f_{GTS}(T) - f_{LRS}(T) = \frac{\rho-\rho^T}{1-\rho^T} - \left(\rho - \frac{1-\rho}{1-\rho^Z}\frac{\rho^{\overline{m}+1} - \rho^T}{1-\rho^{\lambda-1}}\right)$$
$$= \frac{\rho-\rho^T}{1-\rho^T} - \rho + \frac{1-\rho}{1-\rho^Z}\frac{\rho^{\overline{m}+1} - \rho^T}{1-\rho^{\lambda-1}}$$
$$= \frac{\rho-\rho^T-\rho+\rho^{T+1}}{1-\rho^T} + \frac{1-\rho}{1-\rho^Z}\frac{\rho^{\overline{m}+1} - \rho^T}{1-\rho^{\lambda-1}}$$
$$= \frac{-\rho^T(1-\rho)}{1-\rho^T} + \frac{1-\rho}{1-\rho^Z}\frac{\rho^{\overline{m}+1} - \rho^T}{1-\rho^{\lambda-1}}$$

[3]Starting from here and until the end of the proof, in notations $\overline{m}(t)$, $\lambda(t+1)$ argument t means the game duration but not the game stage. The value $\overline{m}(t)$ is defined by formula (5) by substituting $T = t$, and $\lambda(t)$ is also uniquely defined for the game of duration t.

$$= \frac{-\rho^T(1-\rho)}{1-\rho^T} + \frac{1-\rho}{1-\rho^Z}\frac{\rho^{\overline{m}+1}-\rho^T}{1-\rho^{\lambda-1}}$$
$$= (1-\rho)\left(\frac{-\rho^T}{1-\rho^T} + \frac{1}{1-\rho^Z}\frac{\rho^{\overline{m}+1}-\rho^T}{1-\rho^{\lambda-1}}\right).$$

Now we prove that

$$\frac{-\rho^T}{1-\rho^T} + \frac{1}{1-\rho^Z}\frac{\rho^{\overline{m}+1}-\rho^T}{1-\rho^{\lambda-1}} \geq 0. \tag{29}$$

We have:

$$\frac{1}{1-\rho^Z}\frac{\rho^{\overline{m}(T)+1}-\rho^T}{1-\rho^{\lambda(T)-1}} \geq \frac{1}{1-\rho^Z}\frac{\rho^{T-Z}-\rho^T}{1-\rho^{\lambda(T)-1}},$$

since $\overline{m}(T)+1 \leq T-Z$.

Rewriting the last fraction, we obtain:

$$\frac{1}{1-\rho^Z}\frac{\rho^{T-Z}-\rho^T}{1-\rho^{\lambda(T)-1}} = \frac{\rho^{T-Z}}{1-\rho^Z}\frac{1-\rho^Z}{1-\rho^{\lambda(T)-1}} = \frac{\rho^T}{(1-\rho^{\lambda(T)-1})\rho^Z}.$$

Since $\rho^Z = \frac{a-b+\rho(b-d)}{a-d} \leq 1$ and $1-\rho^{\lambda(T)-1} < 1-\rho^T$, we get:

$$\frac{\rho^T}{(1-\rho^{\lambda(T)-1})\rho^Z} \geq \frac{\rho^T}{1-\rho^T}.$$

Using the last inequality, we finally prove inequality (29):

$$\frac{-\rho^T}{1-\rho^T} + \frac{1}{1-\rho^Z}\frac{\rho^{\overline{m}+1}-\rho^T}{1-\rho^{\lambda-1}} \geq \frac{-\rho^T}{1-\rho^T} + \frac{1}{1-\rho^Z}\frac{\rho^{T-Z}-\rho^T}{1-\rho^{\lambda(T)-1}}$$
$$= \frac{-\rho^T}{1-\rho^T} + \frac{\rho^T}{(1-\rho^{\lambda(T)-1})\rho^Z}$$
$$\geq \frac{-\rho^T}{1-\rho^T} + \frac{\rho^T}{1-\rho^T} = 0.$$

This finishes the proof. □

Given the stage payoff matrix, we can find the minimal duration of the game, denoted by $\underline{T}$, such that for any $T \geq \underline{T}$ the minimally acceptable payment scheme exists, but not for $\underline{T}-1$. Let $\underline{T}^{GTS}$ and $\underline{T}^{LRS}$ be the minimal duration of the game under the GT and LR strategies, respectively.

Since $f_{GTS}(t)$ and $f_{LRS}(t)$ are increasing functions of t, the number of stages $\underline{T}^{GTS}$ and $\underline{T}^{LRS}$ are the unique solutions of the corresponding inequalities, respectively:

$$f_{GTS}(\underline{T}-1) < \frac{b-a}{b-d} \leq f_{GTS}(\underline{T}), \tag{30}$$
$$f_{LRS}(\underline{T}-1) < \frac{b-a}{b-d} \leq f_{LRS}(\underline{T}). \tag{31}$$

Corollary 1. The following inequality holds true:

$$\underline{T}^{GTS} \leq \underline{T}^{LRS}.$$

Proof. This inequality follows from Proposition 4. □

The main conclusion is that it is easier to satisfy the conditions for the existence of a MAPS when the players adopt the grim trigger strategy than when they use the more forgiving limited retaliation strategy. Further, the minimal number of stages required is smaller under the GTS. This result is expected since "mild punishment" requires more patience from the players or longer duration of the game. Recall that once we have a minimally acceptable payment scheme, it is possible then to construct an efficient payment scheme.

3.4.2 Undiscounted case

The feasibility condition under LRS is given by

$$\left\lfloor T - 2 - \frac{b-a}{a-d} \right\rfloor \geq 2\frac{b-a}{a-d} + 1.$$

We rewrite this condition in the same form as in the inequalities (25)–(26), i.e., having the same right-hand side expression $\frac{b-a}{b-d}$. After a series of straightforward algebraic manipulations, the above condition becomes

$$1 - \frac{2}{T - 1 - \left\lceil \frac{b-a}{a-d} \right\rceil} \geq \frac{b-a}{b-d}.$$

Let

$$\tilde{f}_{LRS}(T) = 1 - \frac{2}{T - 1 - \left\lceil \frac{b-a}{a-d} \right\rceil} \geq \frac{b-a}{b-d}. \tag{32}$$

To derive the feasibility condition under GTS, recall that in the discounted case we had

$$\frac{\rho - \rho^T}{1 - \rho^T} \geq \frac{b-a}{b-d}.$$

Using l'Hôpital rule, we have

$$\lim_{\rho \to 1} \frac{\rho - \rho^T}{1 - \rho^T} = \lim_{\rho \to 1} \frac{1 - T\rho^{T-1}}{-T\rho^{T-1}} = \frac{T-1}{T}.$$

Consequently, the condition is $\frac{T-1}{T} \geq \frac{b-a}{b-d}$, or equivalently,

$$\tilde{f}_{GTS}(T) = 1 - \frac{1}{T} \geq \frac{b-a}{b-d}. \tag{33}$$

We have the following

Proposition 5. The following holds true:

1. Both $\tilde{f}_{LRS}(T)$ and $\tilde{f}_{GTS}(T)$ are increasing in T.
2. For all parameter values,

$$\lim_{T \to \infty} \tilde{f}_{GTS}(T) = \lim_{T \to \infty} \tilde{f}_{LRS}(T) = 1.$$

3. For all parameter values,

$$\tilde{f}_{GTS}(T) \geq \tilde{f}_{LRS}(T).$$

Proof.

1. This item is obvious.
2. Both limits exist and are equal to 1 taking into account the form of functions $\tilde{f}_{GTS}(T)$ and $\tilde{f}_{LRS}(T)$.
3. To compare $\tilde{f}_{LRS}$ and $\tilde{f}_{GTS}$, compute the difference $\tilde{f}_{GTS} - \tilde{f}_{LRS}$, that is

$$\begin{aligned}\tilde{f}_{GTS}(T) - \tilde{f}_{LRS}(T) &= 1 - \frac{1}{T} - 1 + \frac{2}{T - 1 - \left\lceil \frac{b-a}{a-d} \right\rceil} \\ &= \frac{T + 1 + \left\lceil \frac{b-a}{a-d} \right\rceil}{T(T - 1 - \left\lceil \frac{b-a}{a-d} \right\rceil)} > 0.\end{aligned}$$

We notice that $T - 1 - \left\lceil \frac{b-a}{a-d} \right\rceil > 0$. It follows from inequality (6) assuming that the lower bound is less than the upper bound for at least the first stage $t = 1$.

This finishes the proof. □

In the next corollary, we determine the minimal duration of the game to ensure that the minimally acceptable payment scheme exists. We denote it as $\underline{\tilde{T}}^{GTS}$ and $\underline{\tilde{T}}^{LRS}$ under the GT and LR strategies, respectively.

Corollary 2. The stages $\underline{\tilde{T}}^{GTS}$ and $\underline{\tilde{T}}^{LRS}$ are given by:

$$\underline{\tilde{T}}^{GTS} = \left\lceil \frac{b-d}{a-d} \right\rceil, \tag{34}$$

$$\underline{\tilde{T}}^{LRS} = \left\lceil 1 + \frac{2(b-d)}{a-d} + \left\lceil \frac{b-a}{a-d} \right\rceil \right\rceil. \tag{35}$$

The following inequality holds true:

$$\underline{\tilde{T}}^{GTS} \leq \underline{\tilde{T}}^{LRS}. \tag{36}$$

Proof. Since $\tilde{f}_{GTS}(t)$ and $\tilde{f}_{LRS}(t)$ are increasing functions of t, the formulae (34) and (35) can be easily obtained taking into account inequalities (32) and (33). Inequality (36) immediately follows from Proposition 5. □

Qualitatively speaking, the results are similar in the discounted and undiscounted cases, which is expected.

4 Example

To illustrate our results, we provide a numerical example. Let $T = 15$ and the payoff matrix be given by

	C	N
C	$(7,7)$	$(0,10)$
N	$(10,0)$	$(3,3)$

4.1 Undiscounted case

Players use limited retaliation strategy. Using (6), we obtain $\tilde{\overline{m}} = 12$, which is the maximal retaliation duration at stage 1. Further, the condition for the existence of a MAPS is satisfied. Indeed, we have

$$\tilde{f}_{LRS}(T) = 1 - \frac{2}{T - 1 - \left\lceil \frac{b-a}{a-d} \right\rceil} = 12 \geq \frac{b-a}{b-d} = 2.5. \tag{37}$$

The corresponding payments are given in the second row of Table 1.

Table 1: Payments in minimally acceptable payment schemes under LR $\left(\tilde{\mathcal{P}}^{ma} = \{\tilde{p}_i^{ma}(t)\}\right)$ and GT $\left(\tilde{\mathcal{P}}^{ma,G} = \{\tilde{p}_i^{ma,G}(t)\}\right)$ strategies

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\tilde{p}_i^{ma}(t)$	3	3	3	3	3	3	3	3	3	3	3	**11**	3	3	10
$\tilde{p}_i^{ma,G}(t)$	3	3	3	3	3	3	3	3	3	3	3	**3**	3	3	10

The total payment to each player is 60, and it is lower than the efficient payoff, which is 105 (7×15). The savings after stage $T = 15$ are 45, which is the necessary amount to be allocated along the cooperative trajectory to make the scheme efficient.

Using Proposition 3, we construct an efficient PS that satisfies the MRS property. The corresponding payments are given in the second row of Table 2.

Table 2: Payments in efficient PS under LR strategy $\left(\tilde{\mathcal{P}}^{mrs} = \{\tilde{p}_i^{mrs}(t)\}\right)$ and under GT strategy $\left(\tilde{\mathcal{P}}^{mrs,G} = \{\tilde{p}_i^{mrs,G}(t)\}\right)$

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$\tilde{p}_i^{mrs}(t)$	7	7	7	7	7	7	7	7	7	7	**3**	**11**	7	4	10
$\tilde{p}_i^{mrs,G}(t)$	7	7	7	7	7	7	7	7	7	7	**7**	**7**	7	4	10

Players use grim trigger strategy. Here, the minimally acceptable payments are given in third row of Table 1. The feasibility condition (22) is satisfied, that is, $15 \geq 1.75$. From Table 1, we observe that there is a difference in the payments only at stage 12, where a player gets 11 under LR strategy agreement, while she gets 3 (as in any nonterminal stage) under GT strategy agreement.

Again, summing up the payments we obtain 52, which is less than the 105 that each player would get if (C, C) is played throughout the game. The savings, which are equal to 53, should be allocated along the cooperative trajectory to have an efficient PS. Using Proposition 3, we obtain the efficient payments given in the third row of Table 2. Note that the efficient PS under GT strategy differs from the efficient PS under LR strategy in stages 11 and 12.

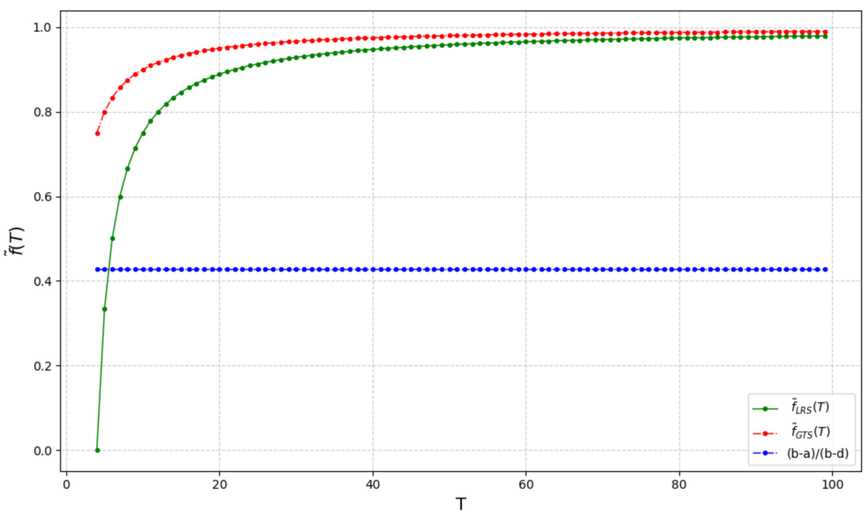

Figure 1: Functions $\tilde{f}_{LRS}$ (green), $\tilde{f}_{GTS}$ (red), and threshold $\frac{b-a}{b-d}$ represented by blue line

Figure 1 graphs the functions $\tilde{f}_{LRS}$ and $\tilde{f}_{GTS}$. Clearly, all statements made in Proposition 5 are correct. The minimal duration of the game under the GT strategy $\underline{\tilde{T}}^{GTS}$ for which there exists a MAPS is 2 stages (given by Equation (34)), and for LR strategy, $\underline{\tilde{T}}^{LRS}$ is 6 stages (given by Equation (35)).

4.2 Discounted case

We consider the same Prisoner's Dilemma game as above, but with a common discount factor $\rho = 0.8$.

Table 3 gives the minimally acceptable payments to any player $i = 1, 2$, under the LR strategy (row 2) and the GT strategy (row 3). Notice that $\overline{m} = 12$ (given by (5)) is the same as in the undiscounted case. Also, as above, the two MAPS differ only in stage 12, with the payment under the LR strategy being larger than under the GT strategy.

Both MAPS are feasible. The inequalities in (26) and (25) read as follows:

$$0.7878 \geq 0.4285 \quad \text{(feasibility condition for LRS)},$$
$$0.7927 \geq 0.4285 \quad \text{(feasibility condition for GTS)}.$$

Table 3: Payments in minimally acceptable payment schemes under LR $\left(\mathcal{P}^{ma} = \{p_i^{ma}(t)\}\right)$ and GT strategy $\left(\mathcal{P}^{ma,G} = \{p_i^{ma,G}(t)\}\right)$

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$p_i^{ma}(t)$	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	**9.008**	4.4	4.4	10
$p_i^{ma,G}(t)$	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	**4.4**	4.4	4.4	10

Since minimally acceptable payment schemes exist and they both are inefficient, we use Proposition 2 to construct efficient payment schemes. The results are in Table 4 in row 2 under the LR strategy, and in row 3 under the GT strategy. As in undiscounted case, the payments are different only in stages 11 and 12. Under GT strategy, the payments in these stages are equal to 7, and it is the same payment as in stages $t = 1, \ldots, 10, 13$. The smallest payment for both PS is in stage 14, when the player is paid less to avoid her deviating before the last stage. She gets 10 instead of 7 in stage 15, where 7 is the stage payoff when (C, C) is played, and 10 is the payoff she can get by individual deviation at the last stage without being punished.

Table 4: Payments in efficient PS under LR strategy $\left(\mathcal{P}^{mrs} = \{p_i^{mrs}(t)\}\right)$ and under GT strategy $\left(\mathcal{P}^{mrs,G} = \{p_i^{mrs,G}(t)\}\right)$

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$p_i^{mrs}(t)$	7	7	7	7	7	7	7	7	7	7	**5.3936**	**9.008**	7	4.6	10
$p_i^{mrs,G}(t)$	7	7	7	7	7	7	7	7	7	7	**7**	**7**	7	4.6	10

Figure 2 graphs the functions $f_{LRS}(T)$ and $f_{GTS}(T)$. Clearly, all the statements made in Proposition 4 are correct.

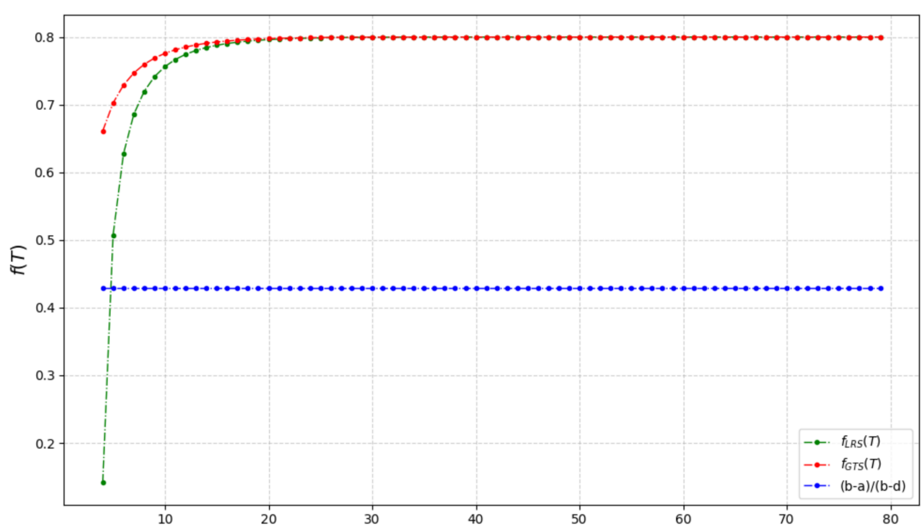

Figure 2: Functions f_{LRS} (green), f_{GTS} (red), and threshold $\frac{b-a}{b-d}$ represented by blue line

Solving inequalities (30) and (31) we find the minimal durations of the game $\underline{T}^{GTS}$ and $\underline{T}^{LRS}$ for which the feasibility property is satisfied under GT and LR strategies, respectively. We obtain $\underline{T}^{GTS} = 2$ from the inequalities $f_{GTS}(1) = 0 < 0.4285 < 0.444 = f_{GTS}(2)$, and $\underline{T}^{LRS} = 5$ stages from $f_{LRS}(4) = 0.1417 < 0.4285 < 0.5074 = f_{LRS}(5)$. Recall that in the undiscounted case, we had $\tilde{\underline{T}}^{GTS} = 2$ (as here) and $\tilde{\underline{T}}^{LRS} = 6$, which is larger. Under the LRS, discounting reduces the number of minimal number of stages needed to satisfy the existence of a MAPS.

5 Conclusions

We proposed a new efficient payment scheme for finitely repeated Prisoner's Dilemma satisfying the minimal required savings property. It ensures that players are paid along the cooperative trajectory in such a way that they keep payoffs for future payments at the minimal level guaranteeing that no player deviates. We examine discounted and undiscounted payoff cases. The influence of duration of the game on the existence of proposed payment schemes is also examined in the paper.

The extension to the case where the two players do not have the same discount rate is of interest. Also, considering a setup more than two players is worth analyzing.

References

Angelova, V., Bruttel, L. V., Güth, W., & Kamecke, U. (2013). Can subgame perfect equilibrium threats foster cooperation? An experimental test of finite-horizon folk theorems. Economic Inquiry, 51, 1345–1356.

Benoit, J., & Krishna, V. (1985) Finitely repeated games. Econometrica, 53(4), 905–922.

Dockner, E., Jørgensen, S., Van Long, N., & Sorger, G. (2000) Differential games in economics and management science. Cambridge University Press.

Dutta, P. (1995). A folk theorem for stochastic games. Journal of Economic Theory, 66, 1–32.

Eswaran, M., & Lewis, T. (1986). Collusive behaviour in finite repeated games with bonding. Economics Letters, 20(3), 213–216.

Flesch, J., & Predtetchinski, A. (2016). On refinements of subgame perfect ε-equilibrium. International Journal of Game Theory, 45(3), 523–542.

Flesch, J., Herings, P.J.-J., Maes, J., & Predtetchinski, A. (2022). Individual upper semicontinuity and subgame perfect ε-equilibria in games with almost perfect information. Economic Theory, 73(2-3), 695–719.

Haurie, A. (1976). A note on nonzero-sum differential games with bargaining solution. Journal of Optimization Theory and Applications, 18(1), 31–39.

Haurie, A., & Pohjola, M. (1987). Efficient Equilibria in a Differential Game of Capitalism. Journal of Economic Dynamics and Control, 11, 65–78.

Haurie, A., & Tolwinski, B. (1985). Definition and properties of cooperative equilibria in a two-player game of infinite duration. Journal of Optimization Theory and Applications, 46(4), 525–534.

Haurie, A., Krawczyk, J. B., & Roche, M. (1994). Monitoring Cooperative Equilibria in a Stochastic Differential Game. Journal of Optimization Theory and Applications, 81, 73–95.

Haurie, A., Krawczyk, J. B., & Zaccour, G. (2012). Games and Dynamic Games. World Scientific Publishing.

Mailath, G.J., Postlewaite, P., & Samuelson, L. (2005). Contemporaneous perfect epsilon-equilibria. Games and Economic Behavior, 53(1), 126–140.

Mailath, G.J., & Samuelson, L. (2006). Repeated Games and Reputations: Long-Run Relationships. Oxford University Press, USA, New York.

Osborne, M., & Rubinstein, A. (1994). A Course in Game Theory. The MIT Press, Cambridge.

Parilina, E.M., Pisareva, A., & Zaccour, G. (2025). Payment schemes for finitely repeated Prisoner's Dilemma games. Theory and Decision. https://doi.org/10.1007/s11238-025-10048-w

Parilina, E., Reddy, P.V., & Zaccour, G. (2022). Theory and Applications of Dynamic Games. A Course on Noncooperative and Cooperative Games Played over Event Trees. Springer Cham.

Parilina, E. M., & Zaccour, G. (2015). Approximated cooperative equilibria for games played over event trees. Operations Research Letters, 43(5), 507–513.

Parilina, E. M., & Zaccour, G. (2022). Payment schemes for sustaining cooperation in dynamic games. Journal of Economic Dynamics and Control, 139, art. no. 104440.

Parilina, E.M., & Zaccour, G. (2024). Payment schemes for sustaining cooperation in dynamic games played over event trees. European Journal of Operational Research, 313(3), 1200–1216.

Parilina, E.M., & Tampieri, A. (2018). Stability and cooperative solution in stochastic games. Theory and Decision, 84 (4), 601–625.

Petrosyan, L. A. (1977). Time consistency of solutions in differential games with many participants. Vestnik Leningradskogo universiteta. Serija 1: matematika, mehanika, astronomija, 19, 46–52. (in Russian)

Petrosyan, L. (2008). Strategically supported cooperation. International Game Theory Review, 10(4), 471–480.

Petrosyan, L. A., & Danilov, N. A. (1979). Time-sonsitent solutions of non-antagonistic differential games with transferable payoffs. Vestnik Leningradskogo universiteta. Serija 1: matematika, mehanika, astronomija, no. 1, 46–54. (in Russian)

Petrosyan, L. A., & Zaccour G. (2018). Cooperative Differential Games with Transferable Payoffs. In: Başar T., Zaccour G. (eds) Handbook of Dynamic Game Theory. Springer, Cham.

Pisareva, A. M., & Parilina, E. M. (2024). Approximate equilibrium in a finitly repeated Prisoner's dilemma. Doklady Mathematics, 110 (suppl. 2), S383–S390.

Radner, R. (1980). Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives. Journal of Economic Theory, 22(2), 136–154.

Tolwinski, B., Haurie A., & Leitmann, G. (1986). Cooperative equilibria in differential games. Journal of Mathematical Analysis and Applications, 119, 182–202.

Yeung, D.W.K., & Petrosyan, L. A. (2012). Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis. Birkhauser Boston, 1–395.

Yeung, D.W.K., & Petrosyan, L. A. (2018). Nontransferable utility cooperative dynamic games, In: Başar T., Zaccour G. (eds) Handbook of Dynamic Game Theory. Springer International Publishing, Cham, 633–670.

Zaccour, G. (2008). Time Consistency in Cooperative Differential Games: A Tutorial. INFOR, 46(1), 81–92.