Vol. 34, No. 2 (2024) | DOI: 10.37190/ord240205

OPEN ACCESS

Operations Research and Decisions

www.ord.pwr.edu.pl

OPERATIONS RESEARCH AND DECISIONS QUARTERLY

On the Nash equilibria of a simple discounted duel

Athanasios Kehagias[1]

[1] *Faculty of Engineering, Aristotle University, Thessaloniki, Greece*
Email address: kehagiat@ece.auth.gr

Abstract

We formulate and study a two-player – duel – game as a nonzero-sum discounted stochastic game. Players $P_1, and P_2$ are standing in place and, in each turn, one or both may shoot at the other player. If P_n shoots at P_m $(m \neq n)$, either he hits and kills him (with probability p_n) or he misses him and P_m is unaffected (with probability $1 - p_n$). The process continues until at least one player dies; if nobody ever dies, the game lasts an infinite number of turns. Each player receives a unit payoff for each turn in which he remains alive; no payoff is assigned to killing the opponent. We show that the always-shooting strategy is a NE but, in addition, the game also possesses so-called cooperative (i.e., non-shooting) Nash equilibria in both stationary and nonstationary strategies. A certain similarity to the repeated Prisoner's Dilemma is also noted and discussed.

Keywords: *game theory, Nash equilibrium, duel*

1. Introduction

In this paper, we study a two-player, duel game played in turns. Players $P_1, and P_2$ are standing in place, and, in each turn, one or both may shoot at the other player. If P_n shoots at P_m $(m \neq n)$, either he hits and kills him or he misses him and P_m is unaffected; the respective probabilities are p_n and $1 - p_n$. The process continues until at least one player dies; it is possible that nobody ever dies and the game lasts an infinite number of turns. We formulate the above as a nonzero-sum discounted stochastic game. The game rules and the player's payoff function will be presented in the next section.

Little work has been done on the duel. As far as we know, it has only been studied as a preliminary step in the study of the truel, in which three stationary players shoot at each other. In early works on the truel [17, 22, 24, 25, 27], the postulated game rules guarantee the existence of exactly one survivor (winner). A more general analysis appears in [23] which considers the possibility of cooperation between the players. This idea is further studied in [19–21, 33]. Recent papers on the truel include [3, 4, 7, 9–11, 13, 30–32] [1].

[1] Let us also note the existence of extensive literature on a quite different type of duel games, which essentially are games of timing [6, 14, 18]. However, this literature is not relevant to the game studied in this paper.

Received 8 May 2023, accepted 12 April 2024, published online 8 July 2024
ISSN 2391-6060 (Online)/

The costs of publishing this issue have been co-financed by the program *Development of Academic Journals* of the Polish Ministry of Education and Science under agreement RCN/SP/0241/2021/1

Many applications of both the duel and the truel have been proposed. A truel model of behavior in a confrontation situation appears in [12] and in politics in [8]. Opinion dissemination has been modeled as a truel in [4]. In [15] business applications have been considered and it is shown that the nuel (an N-person generalization of the duel and truel) the model explains, under appropriate conditions, why weaker companies may strengthen, and why the strongest companies weaken until all companies converge to a common level. The truel has been used in legal studies, as a model of equality issues. Most importantly, the truel has been used to explain the maintenance of biological variation in an ecosystem [5], to model reproduction mechanisms [1] and to explain the existence of "suicidal strategies" used by cells and bacteria [2, 26],

All the above-mentioned works are limited to the analysis of stationary strategies. As shown in the sequel, the duel also possesses Nash equilibria in nonstationary strategy; it is reasonable to assume that this also holds for the truel and the nuel.

While the above papers focus on various forms of the truel and/or nuel, we believe that the duel is interesting in its own right and has not received the attention it deserves. In particular, we will show that, under our formulation, the duel has a certain similarity to the repeated Prisoner's dilemma and possesses left cooperative Nash equilibria in nonstationary strategies.

This paper is structured as follows. In Section 2, we define the game rigorously. In Section 3, we introduce several stationary and nonstationary strategies and compute their expected payoffs. In Section 4, we prove that certain pairs of the previously defined strategies are Nash equilibria. In Section 5, we discuss the obtained results and the connection of the duel to the repeated Prisoner's Dilemma. Finally, in Section 6, we summarize our results and propose some future research directions.

2. The game

In this section, we present a rigorous game theoretic formulation of the duel following [16, 28]. The game involves players P_1, P_2 and proceeds at discrete time steps (rounds) $t \in \{1, 2, \dots\}$. The state at time t is denoted by $s(t)$ and summarizes all relevant information at this time [16]. In the duel game, the state is

$$s(t) = (s_1(t), s_2(t)) \in S = \{(1, 1), (1, 0), (0, 1), (0, 0), (\tau, \tau)\}$$

In more detail, for $n \in \{1, 2\}$, $s_n(t)$ is P_n's state at $t \in \{0, 1, 2, \dots\}$ and can be

$1:$ when P_n is alive,
$0:$ when P_n dies in the current round,
$\tau:$ when one or both players have died in a previous round

In a stochastic game, each player performs an action at each round [16]. In the duel game, P_n's action at $t \in \{1, 2, \dots\}$ is $f_n(t)$, which can be 1 (P_n is shooting) or 0 (P_n is not shooting). If $s_n(t-1) \neq 1$, P_n cannot shoot at t and $f_n(t)$ must equal 0; if $s_n(t-1) = 1$ then $f_n(t)$ can be either 0 or 1. When $f_n(t) = 1$, $s_{-n}(t) = 0$ (i.e., P_{-n} dies[2]) with probability $p_n \in (0, 1)$ and $s_{-n}(t) = 1$ with probability

[2]In the sequel, we use the standard game theoretic notation [28] by which $s_{-1} = s_2$, $s_{-2} = s_1$. The same notation is used for players, actions, etc.

$1 - p_n$. We set $f(t) = (f_1(t), f_2(t))$ and $\mathbf{p} = (p_1, p_2)$. Note that we have assumed that p_1, p_2 are different from both zero and one.

The game starts at an initial state $s(0)$; obviously, the main case of interest is $s(0) = (1,\ 1)$. At times $t \in \{1, 2, \ldots\}$ the players choose simultaneously the actions $f_1(t)$, $f_2(t)$ and the game moves to state $s(t)$ according to the conditional state transition probability $\Pr(s(t)\,|s(t-1),\ f(t))$ [16].

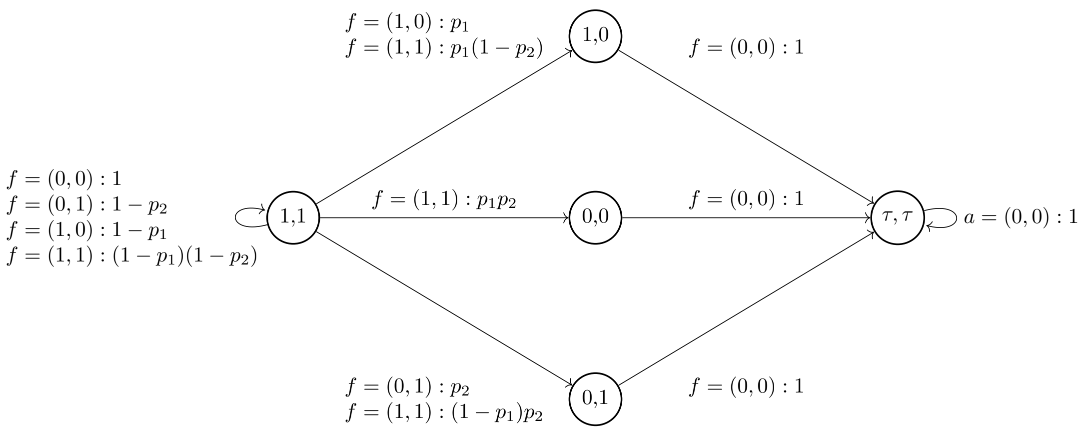

Figure 1. State transition diagram of the duel game.

In Figure 1, we present the state transition diagram, in which the action-dependent transition probabilities are written next to the edges; it is easily verified that these probabilities conform to the game rules. The figure shows that the game starting at $(1,\ 1)$, lasts an infinite number of rounds and two possibilities exist.

1. The game always stays in $(1,\ 1)$ (no player is ever killed).

2. At some t' the game moves to a state $s \in \{(1,\ 0), (0,\ 1), (0,\ 0)\}$ (one or both players are killed) and at $t'+1$ the game moves to the terminal state $(\tau,\ \tau)$, where it stays for ever after.

We use the concept of history presented in [16]. A finite history is a sequence

$$h = s(0)\, f(1)\, s(1) \cdots f(T)\, s(T)$$

an infinite history is an $h = s(0)\, f(1)\, s(1) \cdots$. An admissible history conforms to the game rules; the set of all admissible finite (infinite) histories is denoted by H (H^∞).

We define round and total payoffs as in [16]. Namely, for every history h and for $n \in \{1, 2\}$, we define P_n's total payoff function to be[3]

$$Q_n(h) = \sum_{t=0}^{\infty} \gamma^t q_n(s(t))$$

where $\gamma \in (0,\ 1)$ is the discounting factor and $q_n : S \to \mathbb{R}$ is P_n's round payoff function:

[3] Note that in accordance to [16], we use the unscaled total payoff; but we will also use the "scaled" version (popular with many authors) as explained in Sections 3 and 4.

$$
\begin{aligned}
q_1(\tau,\tau) &= 0, & q_2(\tau,\tau) &= 0\\
q_1(1,1) &= 1, & q_2(1,1) &= 1\\
q_1(1,0) &= \frac{1}{1-\gamma}, & q_2(1,0) &= 0\\
q_1(0,1) &= 0, & q_2(0,1) &= \frac{1}{1-\gamma}\\
q_1(0,0) &= 0, & q_2(0,0) &= 0
\end{aligned}
$$

The above values indicate that each player receives one payoff unit for every turn in which he stays alive; the payoff $q_1(1,0) = \frac{1}{1-\gamma}$ incorporates the infinite payoff sequence $\sum_{t=0}^{\infty} \gamma^t 1 = \frac{1}{1-\gamma}$ (this will result when P_1 kills P_2 and stays alive for an infinite number of subsequent turns). Note that a player receives no direct payoff from killing his opponent, but he has the indirect benefit of removing the possibility of being killed himself.

We use the concept of strategy presented in [16]. A strategy for P_n is a function $\sigma_n : H \to [0,1]$ which corresponds to every finite history h the probability

$$x_n = \sigma_n(h) = \Pr(P_n \text{ shoots at } P_{-n})$$

A stationary strategy (called a Markov strategy in [16]) is a σ_n depending only on the current state s, hence we simply write $x_n = \sigma_n(s)$. A strateg profile is a vector $\sigma = (\sigma_1, \sigma_2)$. We denote the set of all admissible strategies (those which are compatible with the game rules) by Σ and the set of all admissible stationary strategies by $\overline{\Sigma}$.

Given the initial state $s(0)$ and the strategies σ_1 and σ_2, used by P_1 and P_2, respectively, a probability measure is defined on the set of all infinite histories. Since $\gamma \in (0,1)$, the total expected payoffs

$$\forall n \in \{1,2\} : \overline{Q}_n(s(0), \sigma_1, \sigma_2) = \mathbb{E}(Q_n(h) | s(0), \sigma_1, \sigma_2)$$

are well defined [16]. In what follows we will more often use the normalized total expected payoffs, which are defined by

$$\forall n : \widehat{Q}_n(s(0), \sigma_1, \sigma_2) = (1-\gamma)\overline{Q}_n(s(0), \sigma_1, \sigma_2)$$

We have thus formulated the simultaneous duel as a discounted stochastic game, which we will denote by $\Gamma(s(0), \gamma, \mathbf{p})$ or simply $\Gamma(s(0), \gamma)$, when (p_1, p_2) is fixed. Our main interest is in the nonzero-sum game $\Gamma((1,1), \gamma)$. We assume that P_1 and P_2 attempt to reach a Nash equilibrium (NE) as defined in [16, 28], i.e., it is a strategy profile $(\widehat{\sigma}_1, \widehat{\sigma}_2)$ such that

$$\forall n \in \{1,2\} : \forall \sigma_n \in \Sigma : \overline{Q}_n((1,1), \widehat{\sigma}_n, \widehat{\sigma}_{-n}) \geq \overline{Q}_n((1,1), \sigma_n, \widehat{\sigma}_{-n})$$

A refinement of the Nash equilibrium is the subgame perfect equilibrium (SPE). A pair $(\widehat{\sigma}_1, \widehat{\sigma}_2)$ is said to be a SPE iff it is a NE in both the full game and in every subgame [28].

3. Some basic strategies and their payoffs

In this section, we introduce several strategies which will be used in our later exploration of Nash equilibria.

3.1. The stationary strategy σ^S

When P_n ($n \in \{1, 2\}$) uses a stationary admissible strategy σ_n, we have

$$\forall s \in \{(1,\, 0)\,, (0,\, 1)\,, (0,\, 0)\,, (\tau,\, \tau)\} : \sigma_n\,(s) = 0$$

because P_n does not have the option to shoot when either himself or his opponent is dead. It follows that σ_n is fully specified by the value $\sigma_n\,(1,\, 1) = x_n$. Hence we will sometimes write $\overline{Q}_n\,((1,\, 1)\,, x_1,\, x_2)$ in place of $\overline{Q}_n\,((1,\, 1)\,,\, \sigma_1,\, \sigma_2)$.

Let $V_1^S\,(x_1,\, x_2) = \overline{Q}_1\,((1,\, 1)\,, x_1,\, x_2)$; for brevity we will also write simply V_1^S. Then V_1^S satisfies the equation

$$V_1^S = 1 + \gamma p_1 x_1\,(x_2\,(1 - p_2) + (1 - x_2))\,\frac{1}{1-\gamma} + \gamma\,(x_1\,(1 - p_1) + (1 - x_1))\,(x_2\,(1 - p_2) + (1 - x_2))\,V_1^S, \tag{1}$$

obtained by the following reasoning. When the game is in state $s = (1,\, 1)$, P_1's expected payoff is one unit for the current state plus the discounted expected payoff from the subsequent state s', for which we have the following possibilities:

1. $s' = (1,\, 0)$ when P_1 shoots and hits P_2 and P_2 either shoots and misses or does not shoot; the respective probability is $p_1 x_1\,(x_2\,(1 - p_2) + (1 - x_2))$. The total expected payoff of this case is $\overline{Q}_1\,((1,\, 0)\,, x_1,\, x_2) = \dfrac{1}{1-\gamma}$.
2. $s' = (1,\, 1)$ when each of P_1 and P_2 either shoots and misses or does not shoot; the respective probability is $(x_1\,(1 - p_1) + (1 - x_1))\,(x_2\,(1 - p_2) + (1 - x_2))$. In this case we returned to the starting state $(1,\, 1)$ and the additional total expected payoff is again $\overline{Q}_1\,((1,\, 1)\,, x_1,\, x_2)$.
3. We also have the possibilities of moving into $(0,\, 1)$ and $(0,\, 0)$, but these yield zero payoff to P_1, so they are not included in (1).

Solving (1), we get, after some algebraic calculations[4], that

$$V_1^S\,(x_1,\, x_2) = \frac{1 - \gamma\,(1 - p_1 x_1\,(1 - p_2 x_2))}{(1-\gamma)\,(1 - \gamma\,(1 - p_1 x_1)\,(1 - p_2 x_2))} \tag{2}$$

[4]The calculations required to obtain the solution have been performed by the computer algebra system Maple and then verified by hand. This is also true of additional (sometimes quite complicated) calculations required in the rest of the paper.

The corresponding normalized total expected payoff is

$$v_1^S(x_1, x_2) = (1-\gamma)\overline{Q}_1((1,1), x_1, x_2) = \frac{1-\gamma(1-p_1x_1(1-p_2x_2))}{1-\gamma(1-p_1x_1)(1-p_2x_2)} \tag{3}$$

Formulas for $V_2^S(x_1, x_2) = \overline{Q}_2((1,1), x_1, x_2)$ and $v_2^S(x_1, x_2) = (1-\gamma)\overline{Q}_2((1,1), x_1, x_2)$ can be obtained by interchanging the indices 1 and 2 in equations (2), (3).

3.2. The cooperating strategy σ^C

The stationary cooperating (the name will be justified in Section 5) strategy σ^C is defined by

$$\sigma^C(1,1) = 0$$

which means the player never shoots.

Obviously, σ^C is σ^S with $x_n = 0$ and $\overline{Q}_n((1,1), \sigma^C, \sigma^C) = \overline{Q}_n((1,1), 0, 0)$. Hence we obtain $V_1^C = \overline{Q}_1((1,1), \sigma^C, \sigma^C) = \overline{Q}_1((1,1), 0, 0)$ by setting $x_1 = x_2 = 0$ in (2). Consequently, the normalized expected total payoff is

$$v_1^C = (1-\gamma)\overline{Q}_n((1,1), 0, 0) = 1$$

By exchanging the indices 1 and 2 in the above formulas we also get $v_2^c = 1$.

3.3. The defecting strategy σ^D

The stationary defecting (the name will be justified in Section 5) strategy σ^D is defined by

$$\sigma^D(1,1) = 1$$

which means the player always shoots with probability one.

Obviously, σ^D is σ^S with $x_n = 1$ and $\overline{Q}_n((1,1), \sigma^D, \sigma^D) = \overline{Q}_n((1,1), 1, 1)$. Hence we obtain $V_1^D = \overline{Q}_1((1,1), \sigma^D, \sigma^D) = \overline{Q}_1((1,1), 1, 1)$ by setting $x_1 = x_2 = 1$ in (2). We then get the normalized expected total payoff to be

$$v_1^D = (1-\gamma)\overline{Q}_n((1,1), 1, 1) = \frac{1-\gamma(1-p_1(1-p_2))}{1-\gamma(1-p_1)(1-p_2)}$$

and a similar formula for v_2^D.

3.4. The early-shooting strategy $\sigma^{DC(K)}$

The nonstationary early-shooting strategy $\sigma^{DC(K)}$ is to shoot (with probability one) only at times $1, 2, \ldots, K$, where K is a parameter of the strategy. Then we have

$$
\begin{aligned}
&\overline{Q}_1\left((1,\,1)\,,\,\sigma^{DC(K)},\,\sigma^{DC(K)}\right)=1\\
&\quad+\gamma\left(p_1\,(1-p_2)\,\frac{1}{1-\gamma}+(1-p_1)\,(1-p_2)\right)\\
&\quad+\gamma^2\left((1-p_1)\,p_1\,(1-p_2)^2\,\frac{1}{1-\gamma}+(1-p_1)^2\,(1-p_2)^2\right)\\
&\quad\vdots\\
&\quad+\gamma^{K-1}\left((1-p_1)^{K-2}\,p_1\,(1-p_2)^{K-1}\,\frac{1}{1-\gamma}+(1-p_1)^{K-1}\,(1-p_2)^{K-1}\right)\\
&\quad+\gamma^{K}\left((1-p_1)^{K-1}\,p_1\,(1-p_2)^{K}\,\frac{1}{1-\gamma}+(1-p_1)^{K}\,(1-p_2)^{K}\,\overline{Q}_1\left((1,\,1)\,,\,\sigma^{C},\,\sigma^{C}\right)\right)
\end{aligned}
$$

This equation is justified as follows.

1. At time $t=0$, P_1 receives a payoff of one unit.
2. At times $t=1,\,\ldots,K-1$, the expected payoffs are the following:

   a) with probability $(1-p_1)^{t-1}\,p_1\,(1-p_2)^t$, P_1 misses at times $t'=1,\,\ldots,t-1$ and succeeds at time t', while P_2 misses as times $t'=1,\,\ldots,t$. In this case P_1 receives payoff $\frac{1}{1-\gamma}$;

   b) with probability $(1-p_1)^t\,(1-p_2)^t$, both P_1 and P_2 miss at times $t'=1,\,\ldots,t$. In this case P_1 receives payoff 1;

   c) all other possibilities yield zero payoff, so they are not included in the equation.
3. At time $t=K$, the expected payoffs are the following:

   With probability $(1-p_1)^{K-1}\,p_1\,(1-p_2)^K$, P_1 misses at times $t'=1,\,\ldots,K-1$ and succeeds at time $t'=K$, while P_2 misses as times $t'=1,\,\ldots,K$. In this case P_1 receives payoff $\frac{1}{1-\gamma}$.
4. With probability $(1-p_1)^K\,(1-p_2)^K$, both P_1 and P_2 miss at times $t'=1,\,\ldots,K$. In this case both players will never shoot at subsequent times, so we have returned to the starting state $(1,\,1)$ and the additional total expected payoff is $\overline{Q}_1\left((1,\,1)\,,\,\sigma^C,\,\sigma^C\right)$.
5. All other possibilities yield zero payoff, so they are not included in the equation.

   Substituting the expression for $\overline{Q}_1\left((1,\,1)\,,\,\sigma^C,\,\sigma^C\right)$ and after some algebra, we obtain

$$
\begin{aligned}
v_1^{DC(K)}&=(1-\gamma)\,\overline{Q}_1\left((1,\,1)\,,\,\sigma^{DC(K)},\,\sigma^{DC(K)}\right)\\
&=\frac{1+\gamma^{K+1}(1-p_1)^K(1-p_2)^K p_2-\gamma(1-p_1+p_1p_2)}{1-\gamma\,(1-p_1)\,(1-p_2)}
\end{aligned}
$$

and a similar formula for $v_2^{DC(K)}$.

3.5. The late-shooting strategy $\sigma^{CD(K)}$

The nonstationary late-shooting strategy $\sigma^{CD(K)}$ is to shoot (with probability one) only at times $K,\ K+1,\ \ldots$, where K is a parameter of the strategy. Then

$$\overline{Q}_1\left((1,\,1)\,,\,\sigma^{CD(K)},\,\sigma^{CD(K)}\right)=1+\gamma+\cdots+\gamma^{K-1}+ \gamma^{K}\left(p_1\,(1-p_2)\,\frac{1}{1-\gamma}+(1-p_1)\,(1-p_2)\,\overline{Q}_1\left((1,\,1)\,,\,\sigma^{D},\,\sigma^{D}\right)\right)$$

This equation is justified as follows.

1. At times $t=0,\,1,\,\ldots,K-1$, P_1 receives discounted payoff of one unit.
2. At the time $t=K$ we have the following possibilities:

a) with probability $p_1\,(1-p_2)$, P_1 hits P_2, while P_2 misses. In this case P_1 receives payoff $\frac{1}{1-\gamma}$,

b) with probability $(1-p_1)\,(1-p_2)$, both P_1 and P_2 miss. In this case both P_1 and P_2 revert to strategy σ^D and P_1 receives total expected payoff $\overline{Q}_1\left((1,\,1)\,,\,\sigma^{D},\,\sigma^{D}\right)$,

c) all other possibilities yield zero payoffs, so they are not included in the equation.

Substituting in the previously obtained expression for $\overline{Q}_1\left((1,\,1)\,,\,\sigma^{D},\,\sigma^{D}\right)$ we get the following expressions for the normalized expected total payoff:

$$v_1^{CD(K)}=(1-\gamma)\,\overline{Q}_1\left((1,\,1)\,,\,\sigma^{CD(K)},\,\sigma^{CD(K)}\right)=\frac{1-\gamma^{K}p_2-\gamma(1-p_1)\,(1-p_2)}{1-\gamma\,(1-p_1)\,(1-p_2)}$$

and a similar formula for $v_2^{CD(K)}$.

3.6. The periodic-shooting strategy $\sigma^{P(M)}$

The nonstationary periodic-shooting strategy $\sigma^{P(M)}$ is to shoot only at times $M+1, 2M+2,\ \ldots$, where M is a strategy parameter. By reasoning similar to that of the previous cases, we see that

$$\overline{Q}_1\left((1,\,1)\,,\,\sigma^{P(M)},\,\sigma^{P(M)}\right)=1+\gamma+\cdots+\gamma^{M} +\gamma^{M+1}\left(p_1\,(1-p_2)\,\frac{1}{1-\gamma}+(1-p_1)\,(1-p_2)\,\overline{Q}_1\left((1,\,1)\,,\,\sigma^{P(M)},\,\sigma^{P(M)}\right)\right)$$

Solving the above we get the following expression for the normalized expected total payoff:

$$v_1^{P(M)}=(1-\gamma)\,\overline{Q}_1\left((1,\,1)\,,\,\sigma^{P(M)},\,\sigma^{P(M)}\right)=\frac{1-\gamma^{M+1}\,(1-p_1\,(1-p_2))}{1-\gamma^{M+1}\,(1-p_1)\,(1-p_2)}$$

and a similar formula for $v_2^{P(M)}$.

3.7. Grim strategies

We now define grim versions of the previously defined strategies. The first strategy we introduce is the grim-cooperation strategy $\widetilde{\sigma}^{C}$, defined as follows:

$\widetilde{\sigma}^C$: as long as neither player has shot, P_n uses σ^C; if at some turn a player shoots, in all subsequent turns P_n uses σ^D.

Note that while σ^C is stationary, $\widetilde{\sigma}^C$ is, obviously, nonstationary.

We now define grim versions of $\sigma^{DC(K)}$, $\sigma^{CD(K)}$ and $\sigma^{P(M)}$. For these strategies, the condition for reverting to σ_D is slightly (but significantly) different from the one used for $\widetilde{\sigma}^C$. Namely, while in $\widetilde{\sigma}_C$ the condition is that any player shoots, in $\widetilde{\sigma}^{DC(K)}$ the condition is that the other player deviates from $\widetilde{\sigma}^{DC(K)}$ (and similarly for $\widetilde{\sigma}^{P(M)}$ $\sigma^{CD(K)}$ and $\sigma^{P(M)}$).

1. The grim-early-shooting strategy $\widetilde{\sigma}^{DC(K)}$ is defined as follows:

$\widetilde{\sigma}^{DC(K)}$: as long as P_{-n} uses $\sigma^{DC(K)}$, P_n also uses $\sigma^{DC(K)}$;
if at the t-th turn P_{-n} deviates from $\sigma^{DC(K)}$, P_n uses σ^D in all subsequent turns.

2. The grim-late-shooting strategy $\widetilde{\sigma}^{CD(K)}$ is defined as follows:

$\widetilde{\sigma}^{CD(K)}$: as long as P_{-n} uses $\sigma^{CD(K)}$, P_n also uses $\sigma^{CD(K)}$;
if at the t-th turn P_{-n} deviates from $\sigma^{CD(K)}$, P_n uses σ^D in all subsequent turns.

3. The grim-periodic-shooting strategy $\widetilde{\sigma}^{P(M)}$ is defined as follows:

$\widetilde{\sigma}^{P(M)}$: as long as P_{-n} uses $\sigma^{P(M)}$, P_n also uses $\sigma^{P(M)}$;
if at the t-th turn P_{-n} deviates from $\sigma^{P(M)}$, P_n uses σ^D in all subsequent turns.

Again, while the original strategies are stationary, their grim versions are nonstationary.

4. Nash equilibria

We will now present a sequence of propositions; each one indicates that a certain strategy pair is a (stationary or nonstationary) NE and also a SPE; sometimes this will only hold for a certain range of γ and possibly p_1, p_2 values.

Proposition 1. For every $\gamma \in (0, 1)$, the only stationary NE of $\Gamma((1, 1), \gamma)$ are $\left(\sigma^C, \sigma^C\right)$ and $\left(\sigma^D, \sigma^D\right)$. These are also SPE.

Proof. Suppose that P_1 (respectively P_2) uses the stationary strategy σ^S with $\sigma^S(1, 1) = x_1$ (respectively σ^S with $\sigma^S(1, 1) = x_2$). Then P_1's normalized payoff is

$$v_1^S(x_1, x_2) = \frac{1-\gamma(1-p_1x_1(1-p_2x_2))}{1-\gamma(1-p_1x_1)(1-p_2x_2)}$$

Now suppose P_1 switches to some other strategy. Note that we only need to consider stationary strategies. This is a consequence of the following fact, which we will often use in the remainder of the paper. If P_n starts using a stationary strategy σ_n at some time t, then P_{-n}'s best response is also a stationary strategy; because, for a fixed stationary σ_n, P_{-n} has to solve Markov decision process, for which (as is well known) the optimal strategy is stationary [29].

Hence, suppose that P_1 switches to some stationary strategy σ_S, which is fully specified by its value $\sigma_S(1,\,1) = y_1$. Then P_1's normalized payoff becomes

$$v_1^S(y_1, x_2) = \frac{1 - \gamma\left(1 - p_1 y_1\left(1 - p_2 x_2\right)\right)}{1 - \gamma\left(1 - p_1 y_1\right)\left(1 - p_2 x_2\right)}$$

The difference in normalized payoffs is

$$\delta v_1 = \frac{1 - \gamma\left(1 - p_1 x_1\left(1 - p_2 x_2\right)\right)}{1 - \gamma\left(1 - p_1 x_1\right)\left(1 - p_2 x_2\right)} - \frac{1 - \gamma\left(1 - p_1 y_1\left(1 - p_2 x_2\right)\right)}{1 - \gamma\left(1 - p_1 y_1\right)\left(1 - p_2 x_2\right)}$$

$$= \frac{\gamma^2 p_1 p_2 x_2 (x_1 - y_1)(1 - p_2 x_2)}{((1 - \gamma(1 - p_2 x_2)(1 - p_1 y_1))(1 - \gamma(1 - p_2 x_2)(1 - p_1 x_1)))}$$

Now, P_1 has no incentive to switch from x_1 to y_1 iff $\delta v_1 \geq 0$ which is equivalent to

$$\gamma^2 p_1 p_2 x_2 (x_1 - y_1)(1 - p_2 x_2) \geq 0$$

Similarly, P_2 has no incentive to switch from x_2 to y_2 iff

$$\gamma^2 p_1 p_2 x_1 (x_2 - y_2)(1 - p_1 x_1) \geq 0$$

Hence, the following hold for $n \in \{1, 2\}$.

1. If $(x_1,\,x_2) = (0,\,0)$, then P_n has no incentive to change x_n; hence $(x_1,\,x_2) = (0,\,0)$, i.e., $\left(\sigma^C,\,\sigma^C\right)$ is a NE.
2. If $(x_1,\,x_2) = (1,\,1)$, then P_n has no incentive to change x_n; hence $(x_1,\,x_2) = (1,\,1)$, i.e., $\left(\sigma^D,\,\sigma^D\right)$ is a NE.
3. If $(x_1,\,x_2) \in (0,\,1)\,(0,\,1)$, then P_n has incentive to unilaterally change from x_n to 1; hence $(x_1,\,x_2)$ is not a NE.

Suppose that the players use the strategy profile $\left(\sigma^C,\,\sigma^C\right)$ and played t rounds of the duel, reaching state $s(t)$. Now consider any subgame which starts at $s(t)$; in this subgame, because of stationarity of σ^C, the previous history is immaterial and, by the same reasoning as above, $\left(\sigma^C,\,\sigma^C\right)$ is a NE. Hence $\left(\sigma^C,\,\sigma^C\right)$ is a SPE for the full game. By the same reasoning, $\left(\sigma^D,\,\sigma^D\right)$ is also a SPE. This completes the proof. □

Now we will start looking at NE obtained from combinations of grim strategies.

Proposition 2. For every $\gamma \in (0,\,1)$, $\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right)$ is a NE and a SPE of $\Gamma\left((1,\,1), \gamma\right)$.

Proof. Suppose that both P_1 and P_2 use $\widetilde{\sigma}^C$. Then P_1's payoff is

$$\overline{Q}_1\left((1,\,1), \widetilde{\sigma}^C, \widetilde{\sigma}^C\right) = \overline{Q}_1\left((1,\,1)\,\sigma^C,\,\sigma^C\right) = \frac{1}{1-\gamma}$$

Now suppose P_1 deviates from $\widetilde{\sigma}^C$. It suffices to examine the case in which P_1 deviates at $t = 1$; furthermore, after P_1 deviates (i.e., starting at $t = 2$) P_2 will switch to σ^D and P_1 has no incentive to not shoot at any $t \geq 2$ (because, by the same reasoning as in Proposition 1, P_1's best response can be a stationary strategy and then it must be to always shoot with probability one). Hence P_1 is essentially using the strategy $\sigma_1 = \sigma^D$ and his total expected payoff will then be

$$\begin{aligned}
&\overline{Q}_1, \left((1,\, 1)\ \sigma^D, \widetilde{\sigma}^C\right)\\
&= 1 + \gamma\left(p_1 \frac{1}{1-\gamma} + (1-p_1)\left(1 + \gamma \overline{Q}_1\left(\sigma^D,\, \sigma^D\right)\right)\right)\\
&= 1 + \gamma\left(p_1 \frac{1}{1-\gamma} + (1-p_1)\left(1 + \gamma \frac{1-\gamma\left(1-p_1\left(1-p_2\right)\right)}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}\right)\right)\\
&= \frac{1 - p_2(1-p_1)\gamma^3 - (1-p_2)(1-p_1)\gamma}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}
\end{aligned}$$

It follows that the difference in normalized total expected payoffs will be

$$\begin{aligned}
&(1-\gamma)\left(\overline{Q}_1\left((1,\, 1), \widetilde{\sigma}^C, \widetilde{\sigma}^C\right) - \overline{Q}_1\left((1,\, 1)\, \sigma^D, \widetilde{\sigma}^C\right)\right)\\
&= 1 - \frac{1 - p_2(1-p_1)\gamma^3 - (1-p_2)(1-p_1)\gamma}{1-\gamma\left(1-p_1\right)\left(1-p_2\right)}\\
&= \frac{\gamma^3 p_2\left(1-p_1\right)}{1-\gamma\left(1-p_1\right)\left(1-p_2\right)} > 0
\end{aligned}$$

Hence P_1 has no incentive to deviate from $\widetilde{\sigma}^C$. The same can be proved for P_2. Consequently $\left(\widetilde{\sigma}^C,\, \widetilde{\sigma}^C\right)$ is a NE.

To prove that $\left(\widetilde{\sigma}^C,\, \widetilde{\sigma}^C\right)$ is a SPE, we separate all possible subgames into two classes, according to the pre-history (i.e., the history preceding each subgame).

1. Subgames with a pre-history in which neither player has shot. In such subgames, $\widetilde{\sigma}^C$ specifies that each player should not shoot until a shot has been fired in the subgame. In other words, in the subgame the players will use the pair $\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right)$ which, as we have seen is a NE.
2. Subgames with a pre-history in which at least one player has shot. In such subgames, both players should shoot in every riund of the subgame, i.e., $\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right)$ will reduce to $\left(\sigma^D,\, \sigma^D\right)$ which, as we have seen, is also a NE.

Hence $\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right)$ is a SPE. □

In the next proposition, the strategy profile is an NE only for "large enough" γ.

Proposition 3. There exist some $\gamma_0 \in (0,\, 1)$ such that for all $\gamma \in (\gamma_0, 1)$, and for all $K \in \mathbb{N}$, $\left(\widetilde{\sigma}^{DC(K)}, \widetilde{\sigma}^{DC(K)}\right)$ is a NE of $\Gamma\left((1,\, 1), \gamma\right)$.

Proof. Recall that, when both players use $\widetilde{\sigma}^{DC(K)}$, P_1 receives payoff

$$\overline{Q}_1\left((1,\, 1), \widetilde{\sigma}^{DC(K)}, \widetilde{\sigma}^{DC(K)}\right) = \frac{1 + \gamma^{K+1}(1-p_1)^K(1-p_2)^K p_2 - \gamma(1-p_1+p_1p_2)}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}$$

□

Let us show that P_1 has no incentive to use a deviating strategy σ_1.

Case 1. Let us first consider strategies which deviate at times $t \in \{K+1, K+2, \dots\}$, i.e., they shoot after the game has entered the no-shooting phase. For the usual reasons, we only need to consider σ_1 which will shoot at $t = K+1$ and with probability one. In this case

$$\begin{aligned}
\overline{Q}_1\left((1,\,1)\,,\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right) &= A + \gamma^{K+1}\overline{Q}_1\left((1,\,1)\,,\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right)\\
&= A + \gamma^{K+1}\overline{Q}_1\left((1,\,1)\,,\,\sigma^{C},\,\sigma^{C}\right)\\
&= A + \gamma^{K+1}\overline{Q}_1\left((1,\,1)\,,\,\sigma^{C},\,\sigma^{C}\right)\\
\overline{Q}_1\left((1,\,1)\,,\,\sigma^{1},\widetilde{\sigma}^{DC(K)}\right) &= A + \gamma^{K+1}\overline{Q}_1\left((1,\,1)\,,\,\sigma^{1},\widetilde{\sigma}^{DC(K)}\right)\\
&= A + \gamma^{K+1}\overline{Q}_1\left((1,\,1)\,,\,\sigma^{1},\widetilde{\sigma}^{DC(K)}\right)
\end{aligned}$$

where A is the expected payoff summed over times $t \in \{0, \dots, K\}$ and is the same for both strategies used by P_1. Now, for the usual reasons, P_1 will keep shooting at $t \in \{K+2, K+3, \dots\}$ and we will have

$$\begin{aligned}
\overline{Q}_1\left((1,\,1)\,,\,\sigma^{1},\widetilde{\sigma}^{DC(K)}\right) &= p_1\frac{1}{1-\gamma} + (1-p_1)\left(1+\gamma\overline{Q}_1\left((1,\,1)\,,\,\sigma^{D},\,\sigma^{D}\right)\right)\\
&= p_1\frac{1}{1-\gamma} + (1-p_1)\left(1+\gamma\frac{1-\gamma\left(1-p_1\left(1-p_2\right)\right)}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}\right)\\
&= \frac{1-p_2(1-p_1)\gamma^2-(1-p_2)(1-p_1)\gamma}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}
\end{aligned}$$

Then the difference between normalized total expected payoffs is

$$\begin{aligned}
&(1-\gamma)\left(\overline{Q}_1\left((1,\,1)\,,\widetilde{\sigma}^{C},\widetilde{\sigma}^{C}\right) - \overline{Q}_1\left((1,\,1)\,,\,\sigma^{D},\widetilde{\sigma}^{C}\right)\right)\\
&= 1 - \frac{1-p_2(1-p_1)\gamma^2-(1-p_2)(1-p_1)\gamma}{1-\gamma\left(1-p_1\right)\left(1-p_2\right)}\\
&= \frac{p_2(1-p_1)\gamma^2}{1-\gamma\left(1-p_1\right)\left(1-p_2\right)} > 0
\end{aligned}$$

Hence P_1 has no incentive to shoot at $t > K$.

Case 2. Let us consider strategies which deviate at times $t \in \{1, 2, \dots, K\}$, i.e., they do not shoot during the shooting phase. Again, after the first deviation P_1 has no incentive not to shoot. So we only need to consider strategies σ_1 which (a) do not shoot at some $t = L \in \{1, 2, \dots, K\}$ and (b) shoot at all $t \in \{1, 2, \dots, L-1, L+1, \dots\}$. Then, by the usual arguments

$$
\begin{aligned}
&\overline{Q}_1\left((1,\,1)\,,\,\sigma_1,\widetilde{\sigma}^{DC(K)}\right)=1\\
&\quad+\gamma\left(p_1\left(1-p_2\right)\frac{1}{1-\gamma}+\left(1-p_1\right)\left(1-p_2\right)\right)\\
&\quad+\gamma^2\left(\left(1-p_1\right)p_1\left(1-p_2\right)^2\frac{1}{1-\gamma}+\left(1-p_1\right)^2\left(1-p_2\right)^2\right)\\
&\quad\vdots\\
&\quad+\gamma^{L-1}\left(\left(1-p_1\right)^{L-2}p_1\left(1-p_2\right)^{L-1}\frac{1}{1-\gamma}+\left(1-p_1\right)^{L-1}\left(1-p_2\right)^{L-1}\right)\\
&\quad+\gamma^{L}\left(\left(1-p_1\right)^{L-1}p_1\left(1-p_2\right)^{L}\frac{1}{1-\gamma}+\left(1-p_1\right)^{L}\left(1-p_2\right)^{L}\left(1+\gamma V_1^D\right)\right)\\
&\quad=1+\sum_{k=1}^{L-1}\gamma^k\left(\left(1-p_1\right)^{k-1}p_1\left(1-p_2\right)^{k}\frac{1}{1-\gamma}+\left(1-p_1\right)^{k}\left(1-p_2\right)^{k}\right)\\
&\quad+\gamma^{L}\left(1-p_1\right)^{L-1}\left(1-p_2\right)^{L}p_1\frac{1}{1-\gamma}\\
&\quad+\gamma^{L}\left(1-p_1\right)^{L}\left(1-p_2\right)^{L}\left(1+\gamma\frac{1-\gamma\left(1-p_1\left(1-p_2\right)\right)}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}\right)
\end{aligned}
$$

The difference of normalized total expected payoffs is

$$
\delta v_1\left(\gamma\right)=\left(1-\gamma\right)\left(\overline{Q}_1\left((1,\,1)\,,\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right)-\overline{Q}_1\left((1,\,1)\,,\,\sigma_1,\widetilde{\sigma}^{DC(K)}\right)\right)
$$

Note that $\delta v_1\left(\gamma\right)$ is well defined and continuous for all $\gamma\in[0,\,1]$, because the factor $(1-\gamma)$ cancels the $(1-\gamma)$ factor in the denominator of

$$
\overline{Q}_1\left((1,\,1)\,,\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right)-\overline{Q}_1\left((1,\,1)\,,\,\sigma_1,\widetilde{\sigma}^{DC(K)}\right)
$$

After a considerable amount of algebra[5] we find that

$$
\delta v_1\left(1\right)=\frac{\left(1-p_1\right)^K\left(1-p_2\right)^K p_2}{p_1+\left(1-p_1\right)p_2}>0
$$

Since $\delta v_1\left(\gamma\right)$ is continuous, there will exist some $\gamma_0\in(0,\,1)$ such that $\delta v_1\left(\gamma\right)$ will be positive for every $\gamma\in(\gamma_0,1)$ and for every $K\in\mathbb{N}$. Hence, for such values, P_1 has no incentive to deviate during the shooting phase.

[5] Using Maple once again.

Putting together Cases 1 and 2 we see that P_1 has no incentive to deviate from $\widetilde{\sigma}^{DC(K)}$. The same is proved, similarly, for P_2. Hence $\left(\widetilde{\sigma}^{DC(K)}, \widetilde{\sigma}^{DC(K)}\right)$ is a NE.

Next, we present a negative result: mutual late shooting is not a NE.

Proposition 4. For every $\gamma \in (0,\, 1)$ and every $K \in \mathbb{N}$, $\left(\widetilde{\sigma}^{CD(K)}, \widetilde{\sigma}^{CD(K)}\right)$ is not a NE of $\Gamma\left((1,\, 1)\,, \gamma\right)$.

Proof. Recall that

$$\overline{Q}_1\left((1,\, 1)\,,\, \widetilde{\sigma}^{CD(K)},\, \widetilde{\sigma}^{CD(K)}\right) = \frac{1 - \gamma^K p_2 - \gamma(1-p_1)\,(1-p_2)}{(1-\gamma)\,(1-\gamma\,(1-p_1)\,(1-p_2))}$$

We just need to show that P_1 has one profitable deviating strategy σ_1. Let

$$\sigma_1 = \text{do not shoot at } t \in \{1, 2, \ldots, K-2\}\,, \text{shoot at } t \in \{K-1, K, \ldots\}$$

in other words, start shooting one turn before the shooting phase starts. Then, by the usual arguments, P_1's payoff is

$$\overline{Q}_1\left((1,\, 1)\,,\, \sigma_1, \widetilde{\sigma}^{CD(K)}\right) = \sum_{k=0}^{K-2} \gamma^k + \gamma^{K-1}\left(p_1 \frac{1}{1-\gamma} + (1-p_1)\,(1+\gamma V_D)\right)$$

$$= \frac{\left(-\gamma^{K+1} p_2(1-p_1) + 1 - \gamma(1-p_2)(1-p_1)\right)}{(1-\gamma)\,(1-\gamma(1-p_2)(1-p_1))}$$

By appropriate substitutions and algebraic calculations, we get

$$\delta v_1 = (1-\gamma)\left(\overline{Q}_1\left((1,\, 1)\,, \widetilde{\sigma}^{CD(K)}, \widetilde{\sigma}^{CD(K)}\right) - \overline{Q}_1\left((1,\, 1)\,,\, \sigma_1, \widetilde{\sigma}^{CD(K)}\right)\right)$$

$$= -\frac{p_2 \gamma^K\,(1-\gamma\,(1-p_1))}{(1-\gamma(1-p_2)(1-p_1))} < 0$$

Hence P_1 has incentive to switch to σ_1 and $\left(\widetilde{\sigma}^{CD(K)}, \widetilde{\sigma}^{CD(K)}\right)$ is not an NE. □

The next proposition tells us that there exists a set of $(\gamma,\, p_1, p_2)$ combinations for which $\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right)$ is a NE. More specifically, the set of "acceptable" $(\gamma,\, p_1, p_2)$ values is the cube I_M defined below in terms of the parameters γ_M and p_M; the significance of these parameters will be discussed after the proposition is proved.

Proposition 5. For every $M \in \mathbb{N}$, there exists a $\delta_M > 0$ such that if

$$\gamma_M = \frac{9}{10}$$

$$p_M = \frac{1 - \mathrm{e}^{-M}}{10}$$

$$I_M = (\gamma_M - \delta_M, \gamma_M + \delta_M)\,(p_M - \delta_M, p_M + \delta_M)\,(p_M - \delta_M, p_M + \delta_M)\,,$$

then $\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right)$ is a NE of $\Gamma\left((1,\, 1)\,, \gamma, \mathbf{p}\right)$ for every $(\gamma,\, p_1, p_2) \in I_M$.

Proof. Recall that

$$Q_1\left((1,\,1)\,,\widetilde{\sigma}^{P(M)},\,\widetilde{\sigma}^{P(M)}\right) = \frac{1-\gamma^{M+1}\left(1-p_1\left(1-p_2\right)\right))}{\left(1-\gamma\right)\left(1-\gamma^{M+1}\left(1-p_1\right)\left(1-p_2\right)\right)}$$

We will prove that, for every $(\gamma,\,p_1,\,p_2)\in I_M$, P_1 has no incentive to deviate from $\widetilde{\sigma}^{P(M)}$ (the proof forP_2 is identical).

Suppose that P_1 uses some strategy σ_1 by which he shoots at P_2 at some $t\neq i\,(M+1)$. For the usual reasons, it suffices to consider strategies by which P_1 shoots in the first period and with probability one. So suppose that P_1 abstains for all $t\in(1,\,\dots,K)$ and then shoots at P_2 at some $t'=K+1\leq M$. Then the following two possibilities exist.

1. With probability p_1: P_2 is killed and P_1 receives payoff $\dfrac{1}{1-\gamma}$.
2. With probability $1-p_1$: P_2 is missed, P_1 receives payoff one and for all subsequent rounds P_2 will always shoot at P_1 with probability one. In this case, P_1's best response at time $t''>t'$ is to always shoot at P_2 with probability one; hence, starting at the $(t'+1)$-th round, both players use the σ^D strategy. The total expected payoff received by P_1 in this case is $Q_1\left(\sigma^D,\,\sigma^D\right)$.

Hence, assuming P_1 will first shoot at $t=K+1\in\{1,\,\dots,M\}$, by the above reasoning P_1's expected total payoff will be

$$Q_1\left((1,\,1)\,,\,\sigma_1,\widetilde{\sigma}^{P(M)}\right) = \sum_{k=0}^{K}\gamma^k+\gamma^{K+1}\left(p_1\frac{1}{1-\gamma}+\left(1-p_1\right)Q_1\left((1,\,1)\,,\,\sigma^D,\,\sigma^D\right)\right)$$

Substituting the $Q_1\left(\widetilde{\sigma}^{P(M)},\widetilde{\sigma}^{P(M)}\right)$ and $Q_1\left(\sigma^D,\,\sigma^D\right)$ values and performing a considerable amount of algebra we get

$$\begin{aligned}\delta v_1\left(\gamma,\,p_1,p_2\right) &= \left(1-\gamma\right)\left(Q_1\left((1,\,1)\,,\widetilde{\sigma}^{P(M)},\widetilde{\sigma}^{P(M)}\right)-Q_1\left((1,\,1)\,,\,\sigma_1,\widetilde{\sigma}^{P(M)}\right)\ \right)\\ &= \frac{p_2\gamma^{K+2}\left(-\left(1-p_1\right)^2\left(1-p_2\right)\gamma^{M+1}+\left(1-p_1\right)\left(1-p_2\right)\gamma^{M-K}-\gamma^{M-K-1}+1-p_1\right)}{\left(1-\gamma^{M+1}\left(1-p_1\right)\left(1-p_2\right)\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}\end{aligned}$$

Setting $p_1=p_2=p$ we get

$$\delta v_1\left(\gamma,\,p,p\right) = \frac{p\gamma^{K+2}\left(-\left(1-p\right)^3\gamma^{M+1}+\left(1-p\right)^2\gamma^{M-K}-\gamma^{M-K-1}+1-p\right)}{\left(1-\gamma^{M+1}\left(1-p\right)^2\right)\left(1-\gamma\left(1-p\right)^2\right)}$$

The sign of $\delta v_1\left(\gamma,\,p,\,p\right)$ is the same as that of

$$\begin{aligned}f_{M,K}\left(\gamma,\,p\right) &= -\left(1-p\right)^3\gamma^{M+K+3}+\left(1-p\right)^2\gamma^{M+2}-\gamma^{M+1}+\left(1-p\right)\gamma^{K+2}\\ &= f_{1,M,K}\left(\gamma,\,p\right)+f_{2,M,K}\left(\gamma,\,p\right)\end{aligned}$$

with

$$f_{M,K,1}\left(\gamma,\,p\right) = \left(1-p\right)^2\gamma$$

$$^{M+2} - (1-p)^3\,\gamma^{M+K+3} f_{M,K,2}(\gamma,\,p) = -\gamma^{M+1} + (1-p)\,\gamma^{K+2}$$

Now we consider the following cases.

Case 1. $K \le M-2$. Then $M-K-1 \ge 1$. For all M and $K \in \{1,\,\dots,M-2\}$ we have

$$(1-p)^2 > (1-p)^3 \text{ and } \gamma^{M+2} > \gamma^{M+K+3}$$

hence we will always have $f_{1,M,K}(\gamma,\,p) > 0$. To also have $f_{M,K,2}(\gamma,\,p) > 0$ for a specific K, it suffices that

$$\gamma^{M-K-1} < 1-p \Leftrightarrow \gamma < (1-p)^{\frac{1}{M-K-1}}$$

To have $f_{M,K,2}(\gamma,\,p) > 0$ for all $K \in \{1,\,\dots,M-2\}$, it suffices that

$$\gamma < (1-p)^{\frac{1}{M-2}} \tag{4}$$

In other words, for all M we have:

$$\gamma \in \left(0, (1-p)^{\frac{1}{M-2}}\right) \Rightarrow \left(\forall K \in \{1,\,\dots,M-2\} : f_{M,K}(\gamma,\,p) > 0\right)$$
$$\gamma \in \left(0, (1-p)^{\frac{1}{M-2}}\right) \Rightarrow \left(\forall K \in \{1,\,\dots,M-2\} : \delta v_1(\gamma,\,p,p) > 0\right)$$

Case 2. $K = M-1$. In this case

$$\begin{aligned}
&-(1-p)^3\gamma^{M+K+3} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{K+2} > 0\\
&\Leftrightarrow -(1-p)^3\gamma^{M+M-1+3} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{M-1+2} > 0\\
&\Leftrightarrow -(1-p)^3\gamma^{2M+2} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{M+1} > 0\\
&\Leftrightarrow -(1-p)^3\gamma^{M+1} + (1-p)^2\gamma - 1 + (1-p) > 0\\
&\Leftrightarrow -(1-p)^3\gamma^{M+1} + (1-p)^2\gamma - p > 0\\
&\Leftrightarrow -(1-p)^3\gamma^{M+1} + \gamma p^2 - (2\gamma+1)\,p + \gamma > 0
\end{aligned}$$

Let us define the function

$$\overline{f}_M(\gamma,\,p) = -(1-p)^3\gamma^{M+1} + \gamma p^2 - (2\gamma+1)\,p + \gamma$$

By continuity, in a sufficiently small neighborhood of $(\gamma_M, p_M) = \left(\frac{9}{10}, \frac{1-e^{-M}}{10}\right)$, the sign of $\overline{f}_M(\gamma,\,p)$ will be the same as that of

$$h_1(M) = \overline{f}_M(\gamma_M, p_M) = -\frac{\left(9+\mathrm{e}^{-M}\right)^3}{1000}\left(\frac{9}{10}\right)^{M+1} + \frac{9\left(1-\mathrm{e}^{-M}\right)^2}{1000} + \frac{31}{50} + \frac{7\mathrm{e}^{-M}}{25}$$

and it suffices to show that $h_1(M) > 0$ for all M. To this end we first note that

$$h_1(1) = -\frac{\left(9+\mathrm{e}^{-1}\right)^3}{1000}\left(\frac{9}{10}\right)^2 + \frac{9\left(1-\mathrm{e}^{-1}\right)^2}{1000} + \frac{31}{50} + \frac{7\mathrm{e}^{-1}}{25} = 0.60703\cdots > 0$$

Also, letting

$$h_2(M) = -\frac{\left(9+\mathrm{e}^{-M}\right)^3 \left(\frac{9}{10}\right)^{M+1}}{1000} + \frac{31}{50}$$

we have

$$\forall M : h_1(M) > h_2(M)$$

Now, $h_2(M)$ is strictly increasing in M and $h_2(2) = 0.064222\ldots$. Consequently

$$\forall M \geq 2 : h_1(M) > h_2(M) > h_2(2) > 0$$

Hence, finally we have

$$\forall M \geq 1 : \overline{f}_M(\gamma_M, p_M) = h_1(M) > 0.$$

Now, to have

$$\forall M, \forall K \in \{1, \ldots, M-1\} : f_{M,K}(\gamma_M, p_M) > 0$$

we must ensure that (4) holds for $(\gamma, p) = (\gamma_M, p_M)$. In other words, we want $\gamma_M < (1-p_M)^{\frac{1}{M-2}}$ or, equivalently,

$$\frac{9}{10} < \left(1 - \frac{1-e^{-M}}{10}\right)^{\frac{1}{M-2}} = \left(\frac{9}{10} + \frac{e^{-M}}{10}\right)^{\frac{1}{M-2}}.$$

This holds: since for all $M \in \mathbb{N}$ we have $\frac{9}{10} + \frac{e^{-M}}{10} < 1$, we also have

$$\frac{9}{10} < \left(\frac{9}{10} + \frac{e^{-M}}{10}\right) < \left(\frac{9}{10} + \frac{e^{-M}}{10}\right)^{\frac{1}{M-2}}.$$

In short we have shown that

$$\forall M, \forall K \in \{1, \ldots, M-1\} : f_{M,K}(\gamma_M, p_M) > 0$$

$$\forall M, \forall K \in \{1, \ldots, M-1\} : \delta v_1(\gamma_M, p_M, p_M) > 0$$

For all M and K, $\delta v_1(\gamma, p_1, p_2)$ is a continuous function. Hence, for all M, there exists some $\delta_M > 0$ such that

$$\forall K \in \{1, \ldots, M-1\}, \forall (\gamma, p_1, p_2) \in I_M : \delta v_1(\gamma, p_1, p_2) > 0$$

which shows that P_1 has no incentive to deviate from $\widetilde{\sigma}^{P(M)}$. The same argument can be applied to P_2.

Hence, for every $(\gamma, p_1, p_2) \in I_M$, $\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right)$ is a NE of $\Gamma((1,1), \gamma, \mathbf{p})$. □

Remark 1. The fact that $\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right)$ is a NE of $\Gamma((1,1), \gamma, \mathbf{p})$ for every

$$(\gamma, p_1, p_2) \in (\gamma_M - \delta_M, \gamma_M + \delta_M)(p_M - \delta_M, p_M + \delta_M)(p_M - \delta_M, p_M + \delta_M) I_M$$

follows from the continuity of $\delta v_1(\gamma, p_1, p_2)$ and the fact that $\delta v_1(\gamma_M, p_M, p_M) > 0$ or, more explicitly

$$\delta v_1\left(\frac{9}{10}, \frac{1-\mathrm{e}^{-M}}{10}, \frac{1-\mathrm{e}^{-M}}{10}\right) > 0$$

To better understand the significance of the specific values $\gamma_M = \frac{9}{10}$ and $p_M = \frac{1-\mathrm{e}^{-M}}{10}$, the following points should be kept in mind.

1. It is easy to check that $\delta v_1(0, p_1, p_2) = 0$ and $\delta v_1(1, p_1, p_2) < 0$. In other words, to attain a NE some intermediate γ values is required. We used $\gamma_M = \frac{9}{10}$ but (as we have observed by numerical experimentation) there is actually an interval J_M of admissible γ values, with $J_M \subset [0, 1]$. The important thing is that, as proved above, $\gamma_M = \frac{9}{10} \in J_M$ for all M.
2. Once $\gamma_M = \frac{9}{10}$ is fixed, we must determine p_1 and p_2 values which yield a NE. First, we conjectured (after numerical experimentation) that $p_1 = p_2 = p_M = \frac{1-e^{-M}}{10}$ works; and then we proved that it yields a NE. Roughly, the requirement is that the players' marksmanships must be tending to one as M increases to infinity.

5. Some additional remarks

Let us now justify our terms of cooperating and defecting strategy. From the results of Section 3, for $n \in \{1, 2\}$, we have

$$\overline{Q}_n\left((1, 1), \sigma^C, \sigma^C\right) = \frac{1}{1-\gamma}$$

$$\overline{Q}_n\left((1, 1), \sigma^D, \sigma^D\right) = \frac{1-\gamma(1-p_n(1-p_{-n}))}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}$$

It follows that

$$\overline{Q}_n\left((1, 1), \sigma^C, \sigma^C\right) - \overline{Q}_n\left((1, 1), \sigma^D, \sigma^D\right) = \frac{\gamma p_{-n}}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))} > 0$$

In short, just like in PD, it is more profitable for both players not to shoot rather than shoot. Because in our formulation there is no direct profit from killing the opponent, both $\left(\sigma^C, \sigma^C\right)$ and $\left(\sigma^D, \sigma^D\right)$ are NE; however, for both players, $\left(\sigma^C, \sigma^C\right)$ is more profitable NE than $\left(\sigma^D, \sigma^D\right)$. This is the reason for calling $\left(\sigma^C, \sigma^C\right)$ a cooperating, and $\left(\sigma^D, \sigma^D\right)$ a defecting strategy.

All this may be surprising since one would expect that, in a duel, each player's goal will be to eliminate his opponent. It may be supposed that the higher profitability of $\left(\sigma^C, \sigma^C\right)$ follows from our choice of not assigning any direct payoff to killing one's opponent. But this is not true. Even with a positive "killing payoff", $\overline{Q}_n\left((1, 1), \sigma^C, \sigma^C\right)$ can still be greater than $\overline{Q}_n\left((1, 1), \sigma^D, \sigma^D\right)$, provided γ is sufficiently close to one[6]. The reason for the superiority of $\left(\sigma^C, \sigma^C\right)$ is this: if a positive payoff is assigned to survival, this, compounded over an infinite number of turns, can always outweigh the killing payoff. Hence our model can be understood as a more "pacifist" version than the usual duel model[7].

Let us now compare our duel to the PD. In both the PD and the duel, cooperation is more profitable than defection. While $\left(\sigma^C, \sigma^C\right)$ is not a NE in PD, $\left(\sigma^D, \sigma^D\right)$ is an NE in both of them. However, both the

[6] This, as well as additional results regarding the positive killing payoff case will be reported in a future publication.

[7] This point has also been raised by Knuth in the context of the truel [23]. For example, he remarks that *a player who passes is guaranteeing that his opponent has no reason to shoot back, as far as the opponent's survival is concerned.*

duel and the repeated version of PD, possess several NE in grim strategies; the common characteristic of all such equilibria is that they promote cooperation or, in other words, punish defection (shooting). Similarly to the case of repeated PD, it might be possible to prove a "Folk Theorem" for the duel as well; namely that every feasible and individually rational payoff is a NE for γ sufficiently close to one. We intend to study this question in the future.

6. Conclusion

We have formulated the simultaneous shooting duel as a discounted stochastic game and shown that it has two Nash equilibria in stationary strategies, namely the "always-shooting" and the never-shooting strategies; in addition, several nonstationary, cooperation-promoting Nash equilibria also exist. The significance of these results is twofold.

1. In a duel, which at first sight appears to be a purely antagonistic situation, there is scope for the emergence of cooperation; in this connection, the discussed similarity to the repeated PD appears quite relevant.
2. It seems reasonable that, applying similar analysis, we can establish the exi2.stence of nonstationary NE for truels and nuels.

In the future we intend to extend our work in several directions.

First, we want to extend our study and obtain similar results for two variants: (a) the case of non-zero killing payoff and (b) the case of terminal-only payoffs. In addition, we want to formulate and study a version of the duel in which each player wants to kill his opponent in the shortest possible time.

Secondly, we hope to prove a form of Folk theorem, namely that every feasible and individually rational payoff is a NE for γ sufficiently close to one.

Finally, as mentioned above, we want to formulate the nuel as a discounted stochastic game and prove that it possesses NE in nonstationary strategies.

References

[1] Abbott, D. Developments in Parrondo's paradox. In *Applications of Nonlinear Dynamics. Model and Design of Complex Systems* (Berlin, Heidelberg, 2009), V. In, P. Longhini and A. Palacios, Eds., Springer, pp. 307–321.

[2] Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts K., and Walter P. *Molecular Biology of the Cell*. 4th edition, Garland Science, New York, 2002.

[3] Amengual, P. and Toral, R. A Markov chain analysis of truels. In *Proceedings of the 8th Granada Seminar on Computational Physics, Granada, Spain 7–11 February 2005*, Vol. 779 of AIP Conference Proceedings, J. Marro and P. L. Garrido, Eds., (2005), pp. 7–11.

[4] Amengual, P. and Toral, R. Truels, or survival of the weakest. *Computing in Science 8*, 5 (2006), pp. 88–95.

[5] Archetti, M. Survival of the weakest in N-person duels and the maintenance of variation under constant selection. *Evolution 66*, 3 (2012), 637–650.

[6] Barron, E. N. *Game Theory: An Introduction*. John Wiley & Sons, 2013.

[7] Bossert, W., Brams S. J. and Kilgour, D. M. Cooperative vs non-cooperative truels: little agreement, but does that matter? *Games and Economic Behavior 40*, 2 (2002), 185–202.

[8] Brams, S. J. Theory of moves. *American Scientist 81* (1993), 562–570.

[9] Brams, S. J., and Kilgour, D. M. The truel. *Mathematics Magazine 70* (1997), 315-326.

[10] Brams, S. J., and Kilgour, D .M. *Games that End in a Bang or a Whimper*. C. V. Starr Center for Applied Economics, New York University Working Papers, 2001.

[11] Brams, S. J., Kilgour, D. M., and Dawson, B. Truels and the Future. *Math Horizons 10*, 4 (2003), 5–8.

[12] Cole, S. G., Phillips, J. L. and Hartman, E. A. Test of a model of decision processes in an intense conflict situation. *Behavioral Science 22*, 3 (1977), 186–196.

[13] Dorraki, M., Allison, A. and Abbott, D. Truels and strategies for survival. *Scientific Reports 9* (2019), 8996.
[14] Dresher, M. *Games of Strategy: Theory and Applications*. RAND Corporation, 1961.
[15] Dubovik, A. and Parakhonyak, A. *Selective Competition*. Tinbergen Institute Discussion Paper No. 09-072/1, 2009.
[16] Filar, J. and Vrieze, K. *Competitive Markov decision processes*. Springer, New York, 2012.
[17] Gardner, M. *More Mathematical Puzzles and Diversions*, Penguin Books, London, 1966.
[18] Karlin, S. *Mathematical Methods and Theory in Games, Programming, and Economics. The Theory of Infinite Games*. Vol. 2, Addison- Wesley, 1959.
[19] Kilgour, D. M. The simultaneous truel. *International Journal of Game Theory 1* (1971), 229–242.
[20] Kilgour, D. M. The sequential truel. *International Journal of Game Theory 4* (1975), 151-174.
[21] Kilgour, D. M. Equilibrium points of infinite sequential truels. *International Journal of Game Theory 6* (1977), 167–180.
[22] Kinnaird, C. *Encyclopedia of Puzzles and Pastimes*. Citadel Press, 1946.
[23] Knuth, D.E. The triel: A new solution. Journal of Recreational Mathematics 6, 1 (1972), 1–7.
[24] Larsen, H. D., and Moser L. A Dart Game. *American Mathematical Monthly 55*, 10 (1948), 640–641.
[25] Mosteller, F. *Fifty Challenging Problems in Probability with Solutions*. Courier Corporation, 1987.
[26] Ratzke, C., Denk, J., and Gore, J. Ecological suicide in microbes. *Nature Ecology and Evolution 2*, 5 (2018), 867–872.
[27] Shubik, M. Does the fittest necessarily survive? In *Readings in Game Theory and Political Behavior*. Doubleday 1954, pp. 43–47.
[28] Shoham, Y. and Leyton-Brown, K. Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations. Cambridge University Press (2009).
[29] Sobel, M.J. Noncooperative stochastic games. *The Annals of Mathematical Statistics 42*, 6 (1971), 1930–1935.
[30] Toral, R., and Amengual, P. Distribution of winners in truel games. *AIP Conference Proceedings 779*, 1 (2005), 128–141.
[31] Wegener, M. and Mutlu, E. The good, the bad, the well-connected. *International Journal of Game Theory 50*, 3 (2021), 759–771.
[32] Xu, X. Game of the truel. Synthese 185, 1 (supplement) (2012), 19–25.
[33] Zeephongsekul, P. Nash equilibrium points of stochastic N-uels. In *Recent Developments in Mathematical Programming*. CRC Press 1991, pp. 425–452.