Zero-determinant strategies under observation errors in repeated games

Azumi Mamiya[1], Genki Ichinose[1*]

1 Department of Mathematical and Systems Engineering, Shizuoka University, 3-5-1 Johoku, Naka-ku, Hamamatsu, 432-8561, Japan

* ichinose.genki@shizuoka.ac.jp

Abstract

Zero-determinant (ZD) strategies are a novel class of strategies in the Repeated Prisoner's Dilemma (RPD) game discovered by Press and Dyson. This strategy set enforces a linear payoff relationship between a focal player and the opponent regardless of the opponent's strategy. In the RPD game, a discount factor and observation errors are both important because they often happen in society. However, they were not considered in the original discovery of ZD strategies. In some preceding studies, each of them were considered independently. Here, we analytically study the strategies that enforce linear payoff relationships in the RPD game considering both a discount factor and observation errors. As a result, we first revealed that the payoffs of two players can be represented by the form of determinants as shown by Press and Dyson even with the two factors. Then, we searched for all possible strategies that enforce linear payoff relationships and found that both ZD strategies and unconditional strategies are the only strategy sets to satisfy the condition. Moreover, we numerically derived minimum discount rates for the one subset of the ZD strategies in which the extortion factor approaches to infinity. For the ZD strategies whose extortion factor is finite, we numerically derived the minimum extortion factors above which such strategies exist. These results contribute to a deep understanding of ZD strategies in society.

Author summary

Repeated games where two players independently select cooperative or non-cooperative behavior have been used to model interactions of biological organisms. In a real situation, people sometimes cannot observe the direct behaviors that other people select. Instead, they receive signals that reflect other people's behaviors. Those signals are influenced by the environment. Therefore, people sometimes receive the wrong signals. As a result, people mistake other people's behaviors. Hence, in repeated games, assuming such observation errors is important to model biological phenomena close to reality. We mathematically derived that, in the repeated games with observation errors, there are only two types of strategies which enforce a linear payoff relationship to the opponent irrespective of the opponent's strategy. The subsets of the strategies can manipulate the opponent's payoff or enforce an unequal payoff relationship to the opponent. We further numerically revealed the conditions of error rates and a discount factor above which this strategy can exist.

Introduction

Cooperation is a basis for building sustainable societies. In a one-shot interaction, cooperation among individuals is suppressed because cooperation takes costs to the actor while defection does not. This cooperation-defection relationship is well understood by the prisoner's dilemma (PD) game utilized in game theory. In the one-shot PD game, defection is the only Nash equilibrium. When the game is repeated, the situation drastically changes, which is modeled by the repeated prisoner's dilemma (RPD) game [1]. In the RPD game, cooperation will be rewarded by the opponent in the future. In such a situation, cooperation becomes a possible equilibrium. This mechanism is called direct reciprocity [2–4] and makes it possible for players to mutually cooperate in the RPD game.

Evolutionary game theory (EGT) [5] studies how cooperation evolves in the RPD game. Among various cooperative strategies tested in evolutionary games, generous tit-for-tat [6] and win-stay lose-shift [7, 8] were robust to various kinds of evolutionary opponents under noisy conditions. EGT can find strong strategies against various opponents in evolving populations. One missing point was, what is a strong strategy against a direct opponent which utilizes any kind of strategy? In 2012, Press and Dyson suddenly answered this question from a different point of view. Using linear algebraic manipulations, they found a novel class of strategies which contain such ultimate strategies, called zero-determinant (ZD) strategies [9]. ZD strategies impose a linear relationship between the payoffs for a focal player and his opponent regardless of the strategy that the opponent implements. One of the subclasses of ZD strategies is Extortioner which never loses in a one-to-one competition in the RPD game against any opponents.

The discovery of ZD strategies stimulated many researchers. After Stewart and Plotkin raised a question [10], evolution or emergence of ZD strategies became one of the main targets in subsequent studies [11–25]. Then, this research spread in many directions including multiplayer games [19, 26–29], continuous action spaces [28–31], alternating games [31], asymmetric games [32], animal contests [33], human reactions to computerized ZD strategies [34, 35], and human-human experiments [28, 36, 37], which promote an understanding of the nature of human cooperation. For further understanding, see the recent elegant classification of strategies, partners (called "good strategies" in Ref. [11, 38]) and rivals, in direct reciprocity [39]. The utilization of ZD strategies has recently expanded to engineering fields, not just for human cooperation [40–42].

In those ZD studies, no errors were assumed. However, errors (or noise) are unavoidable in human interactions and they may lead to the collapse of cooperation due to negative effects. Thus, the effect of errors has been focused on in the RPD game [43–51]. However, only a few studies have concerned the effect of errors for ZD strategies [52, 53]. There are typically two types of errors: perception errors [45] and implementation errors [46]. Hao et al. [52] and Mamiya and Ichinose [53] considered the former case of the errors where players may misunderstand their opponent's action because the players can only rely on their private monitoring [43, 47] instead of their opponent's direct action. Those studies showed that ZD strategies can exist even in the case that such observation errors are incorporated. In those studies, no discount factor is considered. It is natural to assume that future payoffs will be discounted. Thus, some studies have focused on a discount factor for ZD strategies [30, 31, 54–56] and mathematically found the minimum discount factor above which the ZD strategies can exist [55].

In this study, we search for ZD strategies under the situations that observation errors and a discount factor are both incorporated. We search for the other possible strategies, not just ZD strategies, that enforce a linear payoff relationship between the two players. By formalizing the determinants for the expected payoffs in the RPD game, we mathematically found that only ZD strategies [9] and unconditional strategies [14, 55] are the two types which enforce a linear payoff relationship. We numerically show that the minimum discount factor and extortion rate above which the ZD strategies can exist in the game.

Model

RPD with private monitoring

We consider the symmetric two-person repeated prisoner's dilemma (RPD) game with private monitoring based on the literature [47, 52]. Each player $i \in \{X, Y\}$ chooses an action $a_i \in \{\mathrm{C}, \mathrm{D}\}$ in each round, where C and D imply cooperation and defection, respectively. After the two players conducted the action, player i observes his own action a_i and private signal $\omega_i \in \{g, b\}$ about the opponent's action, where g and b imply *good* and *bad*, respectively. In perfect monitoring, when the opponent takes the action C(D), the focal player always observes the signal $g(b)$. In private monitoring, this is not always true. $\sigma(\boldsymbol{\omega}|\boldsymbol{a})$ is the probability that a signal profile $\boldsymbol{\omega} = (\omega_X, \omega_Y)$ is realized when the action profile is $\boldsymbol{a} = (a_X, a_Y)$ [47]. Let ϵ be the probability that an error occurs to one particular player but not to the other player while ξ be the probability that an error occurs to both players. Then, the probability that an error occurs to neither player is $1 - 2\epsilon - \xi$. For example, when both players take cooperation, $\sigma((g, g)|(\mathrm{C}, \mathrm{C})) = 1 - 2\epsilon - \xi$, $\sigma((g, b)|(\mathrm{C}, \mathrm{C})) = \sigma((b, g)|(\mathrm{C}, \mathrm{C})) = \epsilon$, and $\sigma((b, b)|(\mathrm{C}, \mathrm{C})) = \xi$ are realized.

In each round, player i's realized payoff $u_i(a_i, \omega_i)$ is determined by his own action a_i and signal ω_i, such that $u_i(\mathrm{C}, g) = R$, $u_i(\mathrm{C}, b) = S$, $u_i(\mathrm{D}, g) = T$, and $u_i(\mathrm{D}, b) = P$. Hence, his expected payoff is given by

$$f_i(\boldsymbol{a}) = \sum_{\omega} u_i(a_i, \omega_i)\sigma(\boldsymbol{\omega}|\boldsymbol{a}). \tag{1}$$

The expected payoff is determined by only action profile $\boldsymbol{a}$ regardless of signal profile $\boldsymbol{\omega}$. Thus, the expected payoff matrix is given by

$$\begin{array}{cc} & \begin{array}{cc} \mathrm{C} & \mathrm{D} \end{array} \\ \begin{array}{c} \mathrm{C} \\ \mathrm{D} \end{array} & \begin{pmatrix} R_E & S_E \\ T_E & P_E \end{pmatrix} \end{array}. \tag{2}$$

According to Eq. (1), R_E, S_E, T_E, and P_E are derived as $R_E = R(1 - \epsilon - \xi) + S(\epsilon + \xi)$, $S_E = S(1 - \epsilon - \xi) + R(\epsilon + \xi)$, $T_E = T(1 - \epsilon - \xi) + P(\epsilon + \xi)$, $P_E = P(1 - \epsilon - \xi) + T(\epsilon + \xi)$, respectively. We assume that

$$T_E > R_E > P_E > S_E, \tag{3}$$

and

$$2R_E > T_E + S_E, \tag{4}$$

which dictate the RPD condition with observation errors.

In this paper, we introduce a discount factor to the RPD game with private monitoring. The game is to be played repeatedly over an infinite time horizon but the payoff will be discounted over rounds. Player i's discounted payoff of action profiles $\boldsymbol{a}^t, t \in \{0, 1, ..., \infty\}$ is $w^t f_i(\boldsymbol{a}^t)$ where t is a round. This game can be interpreted as repeated games with a finite but undetermined time horizon. Finally, the average discounted payoff of player i is

$$s_i = (1 - w) \sum_{t=0}^{\infty} w^t f_i(\boldsymbol{a}^t). \tag{5}$$

Determinant form of expected payoff in the RPG game

Here, we proceed to show that Eq. (5) can be represented by a determinant form even for the repeated games with observation errors and a discount factor, as Press and Dyson did for the repeated game without error and no discount factor [9]. The action profiles $\boldsymbol{a}^t$ in Eq. (5) need to be specified to calculate s_i. Those profiles are determined after the strategies of two players are given. Consider player i that adopts memory-one strategies, with which they can use only the outcomes of the last round to decide the action to be submitted in the current round. A memory-one strategy is specified by a 5-tuple; X's strategy is given by a combination of

$$\boldsymbol{p} = (p_1, p_2, p_3, p_4; p_0), \tag{6}$$

where $0 \le p_j \le 1, j \in \{0, 1, 2, 3, 4\}$. The subscripts 1, 2, 3, and 4 of p mean previous outcome Cg, Cb, Dg and Db, respectively. In Eq. (6), p_1 is the conditional probability that X cooperates when X cooperated and observed signal g in the last round, p_2 is the conditional probability that X cooperates when X cooperated and observed signal b in the last round, p_3 is the conditional probability that X cooperates when X defected and observed signal g in the last round, and p_4 is the conditional probability that X cooperates when X defected and observed signal b in the last round. Finally, p_0 is the probability that X cooperates in the first round. Similarly, Y's strategy is specified by a combination of

$$\boldsymbol{q} = (q_1, q_2, q_3, q_4; q_0), \tag{7}$$

where $0 \le q_j \le 1, j \in \{0, 1, 2, 3, 4\}$.

Define $\boldsymbol{v}(t) = (v_1(t), v_2(t), v_3(t), v_4(t))$ as the stochastic state of two players in round t where the subscripts 1, 2, 3, and 4 of v imply the stochastic states (C,C), (C,D), (D,C), and (D,D), respectively. $v_1(t)$ is the probability that both players cooperate in round t, $v_2(t)$ is the probability that X cooperates and Y defects in round t, and so forth. Then, the expected payoff to player X in round t is given by $\boldsymbol{v}(t)\boldsymbol{S}_X$, where $\boldsymbol{S}_X^T = (R_E, S_E, T_E, P_E)$. The expected per-round payoff to player X in the repeated game is given by

$$s_X = (1-w)\sum_{t=0}^{\infty} w^t \boldsymbol{v}(t)\boldsymbol{S}_X, \tag{8}$$

where $0 < w < 1$. The initial stochastic state is given by

$$\boldsymbol{v}(0) = (p_0 q_0, p_0(1-q_0), (1-p_0)q_0, (1-p_0)(1-q_0)). \tag{9}$$

The state transition matrix M of these repeated games with observation errors is given by

$$M = \begin{pmatrix}
\begin{pmatrix} \tau p_1 q_1 \\ +\epsilon p_1 q_2 \\ +\epsilon p_2 q_1 \\ +\xi p_2 q_2 \end{pmatrix} &
\begin{pmatrix} \tau p_1(1-q_1) \\ +\epsilon p_1(1-q_2) \\ +\epsilon p_2(1-q_1) \\ +\xi p_2(1-q_2) \end{pmatrix} &
\begin{pmatrix} \tau(1-p_1)q_1 \\ +\epsilon(1-p_1)q_2 \\ +\epsilon(1-p_2)q_1 \\ +\xi(1-p_2)q_2 \end{pmatrix} &
\begin{pmatrix} \tau(1-p_1)(1-q_1) \\ +\epsilon(1-p_1)(1-q_2) \\ +\epsilon(1-p_2)(1-q_1) \\ +\xi(1-p_2)(1-q_2) \end{pmatrix} \\
\begin{pmatrix} \epsilon p_1 q_3 \\ +\xi p_1 q_4 \\ +\tau p_2 q_3 \\ +\epsilon p_2 q_4 \end{pmatrix} &
\begin{pmatrix} \epsilon p_1(1-q_3) \\ +\xi p_1(1-q_4) \\ +\tau p_2(1-q_3) \\ +\epsilon p_2(1-q_4) \end{pmatrix} &
\begin{pmatrix} \epsilon(1-p_1)q_3 \\ +\xi(1-p_1)q_4 \\ +\tau(1-p_2)q_3 \\ +\epsilon(1-p_2)q_4 \end{pmatrix} &
\begin{pmatrix} \epsilon(1-p_1)(1-q_3) \\ +\xi(1-p_1)(1-q_4) \\ +\tau(1-p_2)(1-q_3) \\ +\epsilon(1-p_2)(1-q_4) \end{pmatrix} \\
\begin{pmatrix} \epsilon p_3 q_1 \\ +\tau p_3 q_2 \\ +\xi p_4 q_1 \\ +\epsilon p_4 q_2 \end{pmatrix} &
\begin{pmatrix} \epsilon p_3(1-q_1) \\ +\tau p_3(1-q_2) \\ +\xi p_4(1-q_1) \\ +\epsilon p_4(1-q_2) \end{pmatrix} &
\begin{pmatrix} \epsilon(1-p_3)q_1 \\ +\tau(1-p_3)q_2 \\ +\xi(1-p_4)q_1 \\ +\epsilon(1-p_4)q_2 \end{pmatrix} &
\begin{pmatrix} \epsilon(1-p_3)(1-q_1) \\ +\tau(1-p_3)(1-q_2) \\ +\xi(1-p_4)(1-q_1) \\ +\epsilon(1-p_4)(1-q_2) \end{pmatrix} \\
\begin{pmatrix} \xi p_3 q_3 \\ +\epsilon p_3 q_4 \\ +\epsilon p_4 q_3 \\ +\tau p_4 q_4 \end{pmatrix} &
\begin{pmatrix} \xi p_3(1-q_3) \\ +\epsilon p_3(1-q_4) \\ +\epsilon p_4(1-q_3) \\ +\tau p_4(1-q_4) \end{pmatrix} &
\begin{pmatrix} \xi(1-p_3)q_3 \\ +\epsilon(1-p_3)q_4 \\ +\epsilon(1-p_4)q_3 \\ +\tau(1-p_4)q_4 \end{pmatrix} &
\begin{pmatrix} \xi(1-p_3)(1-q_3) \\ +\epsilon(1-p_3)(1-q_4) \\ +\epsilon(1-p_4)(1-q_3) \\ +\tau(1-p_4)(1-q_4) \end{pmatrix}
\end{pmatrix}, \tag{10}$$

where $\tau = 1 - 2\epsilon - \xi$. Then, we obtain

$$\boldsymbol{v}(t) = \boldsymbol{v}(0) M^t. \tag{11}$$

By substituting Eq. (11) in Eq (8), we obtain

$$\begin{aligned} s_X &= (1-w)\boldsymbol{v}(0) \sum_{t=0}^{\infty} (wM)^t \boldsymbol{S}_X \\ &= (1-w)\boldsymbol{v}(0)(I - wM)^{-1} \boldsymbol{S}_X, \end{aligned} \tag{12}$$

where I is the 4×4 identity matrix. Then, let

$$\boldsymbol{v}^T = (v_1, v_2, v_3, v_4) = (1-w)\boldsymbol{v}(0)(I - wM)^{-1} \tag{13}$$

be the mean distribution of $\boldsymbol{v}(t)$. Additionally, we define

$$M_0 = \begin{pmatrix} p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0 q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \end{pmatrix}. \tag{14}$$

Because $v_1 + v_2 + v_3 + v_4 = 1$ (S1 Appendix), the following holds (S2 Appendix)

$$\boldsymbol{v}(0) = \boldsymbol{v}^T M_0. \tag{15}$$

By substituting Eq. (15) in Eq. (13) and multiplying both sides of the equation by $(I - wM)$ from the right, we obtain

$$\boldsymbol{v}^T (I - wM) = (1-w)\boldsymbol{v}^T M_0. \tag{16}$$

Equation (16) and $M' \equiv wM + (1-w)M_0 - I$ yield

$$\boldsymbol{v}^T M' = 0. \tag{17}$$

Applying Cramer's rule to matrix M', we obtain

$$Adj(M')M' = 0, \tag{18}$$

where $Adj(M')$ is the adjugate matrix of M'. Eqs. (17) and (18) imply that every row of $Adj(M')$ is proportional to $\boldsymbol{v}$. Therefore, $\boldsymbol{v}$ is solely represented by the components of matrix M'. Choosing the fourth row of the matrix $Adj(M')$, we see that $\boldsymbol{v}$ is composed of the determinant of the 3×3 matrixes formed from the first three columns of M'. We add the first column of M' into the second and third columns. Even by this manipulation, this determinant is unchanged. The result of these manipulations is a formula for the dot product of an arbitrary vector $\boldsymbol{f}^T = (f_1, f_2, f_3, f_4)$ with the fourth column vector $\boldsymbol{u}$ of matrix M', which can be represented by the form of the determinant

$$\begin{aligned} &\boldsymbol{u} \cdot \boldsymbol{f} = \\ &\begin{vmatrix} w(\tau p_1 q_1 + \epsilon p_1 q_2 + \epsilon p_2 q_1 + \xi p_2 q_2) - 1 + p_0 q_0(1-w) & w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) & w(\mu q_1 + \eta q_2) - 1 + q_0(1-w) & f_1 \\ w(\epsilon p_1 q_3 + \xi p_1 q_4 + \tau p_2 q_3 + \epsilon p_2 q_4) + p_0 q_0(1-w) & w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) & w(\mu q_3 + \eta q_4) + q_0(1-w) & f_2 \\ w(\epsilon p_3 q_1 + \tau p_3 q_2 + \xi p_4 q_1 + \epsilon p_4 q_2) + p_0 q_0(1-w) & w(\mu p_3 + \eta p_4) + p_0(1-w) & w(\eta q_1 + \mu q_2) - 1 + q_0(1-w) & f_3 \\ w(\xi p_3 q_3 + \epsilon p_3 q_4 + \epsilon p_4 q_3 + \tau p_4 q_4) + p_0 q_0(1-w) & w(\eta p_3 + \mu p_4) + p_0(1-w) & w(\eta q_3 + \mu q_4) + q_0(1-w) & f_4 \end{vmatrix} \\ &\equiv D(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{f}) \end{aligned} \tag{19}$$

where $\mu = 1 - \epsilon - \xi$ and $\eta = \epsilon + \xi$. Furthermore, Eq. (19) should be normalized to have its components sum to 1 by $\boldsymbol{u} \cdot \mathbf{1}$, where $\mathbf{1} = (1, 1, 1, 1)$. Then, we obtain the dot product of an arbitrary vector $\boldsymbol{f}$ with mean distribution $\boldsymbol{v}$. Replacing the last column of $D(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{f})$ with player X's and Y's expected payoff vector, respectively, we obtain their per-round expected payoffs:

$$s_X = \boldsymbol{v} \cdot \boldsymbol{S}_X = \frac{\boldsymbol{u} \cdot \boldsymbol{S}_X}{\boldsymbol{u} \cdot \mathbf{1}} = \frac{D(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{S}_X)}{D(\boldsymbol{p}, \boldsymbol{q}, \mathbf{1})}, \tag{20}$$

$$s_Y = \boldsymbol{v} \cdot \boldsymbol{S}_Y = \frac{\boldsymbol{u} \cdot \boldsymbol{S}_Y}{\boldsymbol{u} \cdot \mathbf{1}} = \frac{D(\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{S}_Y)}{D(\boldsymbol{p}, \boldsymbol{q}, \mathbf{1})}. \tag{21}$$

When we set $w = 1$, Eq. (19) corresponds to Eq. (2) of [52]. By using Eq. (19), we can calculate players' per-round expected payoffs when $0 \leq w \leq 1$ by the form of the determinants. $w = 0$ can be interpreted as a one-shot game and $w = 1$ is the case where future payoffs are not discounted.

Results

Mathematical analysis

Since we are interested in the payoff relationship between the two players, we linearly combine those payoffs represented by Eqs. (20) and (21). The linear combination of s_X and s_Y can also be represented by the form of the determinant:

$$\alpha s_X + \beta s_Y + \gamma = \frac{D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \mathbf{1})}{D(\boldsymbol{p}, \boldsymbol{q}, \mathbf{1})}, \tag{22}$$

where α, β, and γ, are arbitrary constants. The numerator of the right side of Eq. (22) is expressed in the following:

$$D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \mathbf{1}) =$$
$$\begin{vmatrix} w(\tau p_1 q_1 + \epsilon p_1 q_2 + \epsilon p_2 q_1 + \xi p_2 q_2) - 1 + p_0 q_0 (1-w) & w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) & w(\mu q_1 + \eta q_2) - 1 + q_0(1-w) & \alpha R_E + \beta R_E + \gamma \\ w(\epsilon p_1 q_3 + \xi p_1 q_4 + \tau p_2 q_3 + \epsilon p_2 q_4) + p_0 q_0 (1-w) & w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) & w(\mu q_3 + \eta q_4) + q_0(1-w) & \alpha S_E + \beta T_E + \gamma \\ w(\epsilon p_3 q_1 + \tau p_3 q_2 + \xi p_4 q_1 + \epsilon p_4 q_2) + p_0 q_0 (1-w) & w(\mu p_3 + \eta p_4) + p_0(1-w) & w(\eta q_1 + \mu q_2) - 1 + q_0(1-w) & \alpha T_E + \beta S_E + \gamma \\ w(\xi p_3 q_3 + \epsilon p_3 q_4 + \epsilon p_4 q_3 + \tau p_4 q_4) + p_0 q_0 (1-w) & w(\eta p_3 + \mu p_4) + p_0(1-w) & w(\eta q_3 + \mu q_4) + q_0(1-w) & \alpha P_E + \beta P_E + \gamma \end{vmatrix}, \tag{23}$$

If Eq. (23) is zero, the relationship between the two players' payoffs becomes linear

$$\alpha s_X + \beta s_Y + \gamma = 0. \tag{24}$$

Thus, we search for all of the solutions such that $D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \mathbf{1}) = 0$.

Press and Dyson [9] (without error) and Hao et al. [52] (with observation errors) searched for the case that second and fourth columns of the determinant take the same value. This makes the determinant zero. Also, Mamiya and Ichinose [53] searched for all the cases, from all possibilities, that make the determinant zero with observation errors. Here, we extend Mamiya and Ichinose [53] to the case with both observation errors and a discount factor.

The following determinant theorem gives such a condition.

Theorem 1 For $n \times n$ matrix A, the following holds:

$$det(\mathrm{A}) = 0 \quad \Leftrightarrow \quad \text{The columns of matrix A are linearly dependent vectors.}$$

We define $\boldsymbol{d}_i, i \in \{1,2,3,4\}$ as i-th column vector of the determinant of Eq. (23). From the above theorem, if the columns of the determinant of Eq. (23) are linearly dependent vectors, there exist real numbers $s, t, u, v, \alpha, \beta$, and γ, except for the trivial solution $((s,t,u,v) = (0,0,0,0)$,$(\alpha,\beta,\gamma) = (0,0,0))$, such that

$$s\boldsymbol{d}_1 + t\boldsymbol{d}_2 + u\boldsymbol{d}_3 + v\boldsymbol{d}_4 = \boldsymbol{0}, \tag{25}$$

where vector $\boldsymbol{0}$ denotes a zero vector. We give the detailed calculation in S3 Appendix.

As a result, we found that, in the RPD game even with observation errors (imperfect monitoring) and a discount factor, the only strategies that impose a linear payoff relationship between the two players' payoffs are either

$$\begin{aligned} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) &= \alpha R_E + \beta R_E + \gamma \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) &= \alpha S_E + \beta T_E + \gamma \\ w(\mu p_3 + \eta p_4) + p_0(1-w) &= \alpha T_E + \beta S_E + \gamma \\ w(\eta p_3 + \mu p_4) + p_0(1-w) &= \alpha P_E + \beta P_E + \gamma \end{aligned} \tag{26}$$

or

$$p_0 = p_1 = p_2 = p_3 = p_4. \tag{27}$$

Existence of subsets of ZD strategies

Since observation errors and a discount factor are considered, in general, the ranges in which ZD strategies can exist are narrowed. Ichinose and Masuda mathematically showed the minimum discount rates above which Equalizer (a subclass of ZD strategies) can exist [55]. Here, we numerically address threshold values where subsets of ZD strategies can exist.

Minimum discount factor for Equalizer

Equalizer strategies are a subclass of ZD strategies. We first show minimum discount factor w_c for Equalizer when observation errors ϵ and ξ are given. Equalizer can fix the opponent payoff no matter what the opponent takes, which means that

$$\beta s_Y + \gamma = 0. \tag{28}$$

This is obtained by substituting $\alpha = 0$ into Eq. (24). We substitute $\alpha = 0$ into Eq. (26) to obtain Equalizer

$$\begin{aligned} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) &= \beta R_E + \gamma \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) &= \beta T_E + \gamma \\ w(\mu p_3 + \eta p_4) + p_0(1-w) &= \beta S_E + \gamma \\ w(\eta p_3 + \mu p_4) + p_0(1-w) &= \beta P_E + \gamma. \end{aligned} \tag{29}$$

If we solve Eq. (29) for β, γ, p_2 and p_3,

$$\begin{aligned} \beta &= -\frac{(1 - wp_1 + wp_4)(\mu - \eta)}{\mu(R_E - P_E) - \eta(T_E - S_E)} \\ \gamma &= \frac{(1 - wp_1 - p_0 + wp_0)(\mu P_E - \eta S_E) + (p_0 - wp_0 + wp_4)(\mu R_E - \eta T_E)}{\mu(R_E - P_E) - \eta(T_E - S_E)} \\ p_2 &= \frac{p_1(\mu(T_E - P_E) - \eta(R_E - S_E)) - (\frac{1}{w} + p_4)(T_E - R_E)}{\mu(R_E - P_E) - \eta(T_E - S_E)} \\ p_3 &= \frac{(\frac{1}{w} - p_1)(P_E - S_E) + p_4(\mu(R_E - S_E) - \eta(T_E - P_E))}{\mu(R_E - P_E) - \eta(T_E - S_E)} \end{aligned} \tag{30}$$

are obtained. By substituting β and γ into Eq. (28), player Y's payoff is fixed at

$$s_Y = \frac{(1 - wp_1 - p_0 + wp_0)(\mu P_E - \eta S_E) + (p_0 - wp_0 + wp_4)(\mu R_E - \eta T_E)}{(1 - wp_1 + wp_4)(\mu - \eta)}. \quad (31)$$

Equations (30) and (31) correspond to Eq. (10) in [52] when $w = 1$.

Equalizer must satisfy the condition $0 \leq p_i \leq 1$ in Eq. (29). The existence of Equalizer strategies also depends on w, ϵ and ξ. We numerically find the minimum discount rate w_c and the condition of (ϵ, ξ) that Equalizer exists. $w \geq w_c$ is the condition for w under which Equalizer strategies exist.

Figure 1 shows w_c when $\epsilon + \xi$ is given. We set $(T, R, P, S) = (1.5, 1, 0, -0.5)$ and excluded the case $\epsilon + \xi > 1/3$ because $T_E > R_E > P_E > S_E$ is not satisfied under the situation. Note that the effects of ϵ and ξ are the same because $\eta = \epsilon + \xi$ and $\mu = 1 - \epsilon - \xi$ in Eq. (29) includes both ϵ and ξ. When there was no error ($\epsilon + \xi = 0$), w_c was about 0.33. When the errors were $\epsilon + \xi = 0.1$ and 0.2, w_c were about 0.52 and 0.93. As a result, we found that $w \geq w_c$ for Equalizer becomes larger as the error is increased.

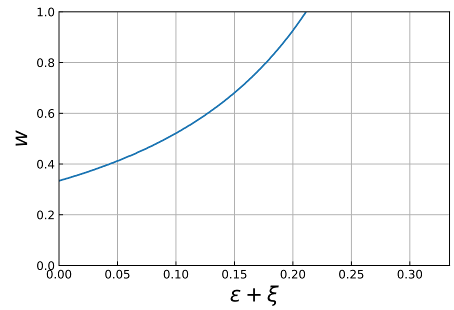

Fig 1. Minimum discount rate w_c for Equalizer.

Minimum extortion factor for ZD strategies with $1 \leq \chi < \infty$

Next we focus on other types of ZD strategies which include Extortion [9] and Generous [23]. In Eq. (26), we substitute $\alpha = \phi\chi, \beta = -\phi$, and $\gamma = \phi(1 - \chi)\kappa$ to obtain

$$\begin{aligned} w(\mu p_1 + \eta p_2) - 1 + p_0(1 - w) &= \phi[(R_E - \kappa) - \chi(R_E - \kappa)] \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1 - w) &= \phi[(S_E - \kappa) - \chi(T_E - \kappa)] \\ w(\mu p_3 + \eta p_4) + p_0(1 - w) &= \phi[(T_E - \kappa) - \chi(S_E - \kappa)] \\ w(\eta p_3 + \mu p_4) + p_0(1 - w) &= \phi[(P_E - \kappa) - \chi(P_E - \kappa)]. \end{aligned} \quad (32)$$

In Eq. (32), we obtain Extortion when $\kappa = P$ and Generous when $\kappa = R$ with $1 \leq \chi < \infty$ when there are no errors ($\epsilon + \xi = 0$) and no discount factor ($w = 1$). Note that $\chi \to \infty$ in Eq. (32) corresponds to Equalizer [52]. When there are no errors $(\epsilon, \xi) = (0, 0)$ and no discount factor ($w = 1$), Eq. (32) corresponds to Extortion in Press and Dyson [9] when $\kappa = P$ and Generous in Stewart and Plotkin [23] when $\kappa = R$.

We numerically calculated the minimum extortion factor χ_c for subsets of ZD strategies with $1 \leq \chi < \infty$ to exist (Fig. 2). Each curve corresponds to each w as shown in the legend. The area surrounded by each curve and the vertical axis is the region of χ which can be utilized by the ZD strategies when $\epsilon + \xi$ is fixed. As the error $\epsilon + \xi$ becomes larger and the discount factor w becomes smaller, the minimum extortion factor χ_c becomes larger.

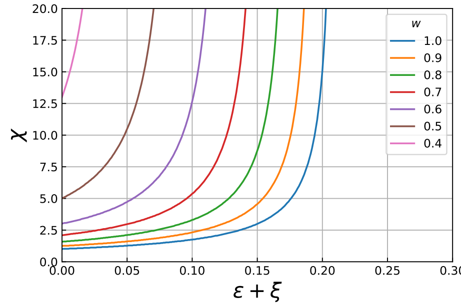

Fig 2. Minimum extortion factor χ_c for subsets of ZD strategies with $1 \leq \chi < \infty$. $(T, R, P, S) = (1.5, 1, 0, -0.5)$. We adopted p_0 and κ so that χ_c was minimized.

Numerical examples of representative ZD and unconditional strategies under errors in repeated games

We numerically demonstrate that ZD and unconditional strategies can impose a linear relationship between the two players' payoffs while others cannot in the RPD game under errors. We take up Extortion and Equalizer as the representative of ZD strategies, ALLD as the representative of unconditional strategies, and Win-Stay-Lose-Shift (WSLS) as neither ZD nor unconditional strategies.

Figure 3 shows the relationship between the two players' expected payoffs per game with payoff vector $(T, R, P, S) = (1.5, 1, 0, -0.5)$. The gray quadrangle in each panel represents the feasible set of payoffs. We fixed one particular strategy for player X (vertical line) and randomly generated 1,000 strategies that satisfy $0 \leq q_0, q_1, q_2, q_3, q_4 \leq 1$ for player Y (horizontal axis). Thus, each black dot represents the payoff relationship between two players. In addition, the blue and red are the particular cases for player Y. Red is the case that player Y is ALLD and blue is the case that player Y is ALLC. We set $w = 1$ for Figs. 3A–D and $w = 0.9$ for Figs. 3E–H. In each figure, we used three error rates $\epsilon + \xi = 0, 0.1$ and 0.2.

Figures 3A and E show the case with a WSLS strategy vs. $1000 + 2$ strategies. In the case, $\xi = 0$ is fixed and ϵ is caried to 0, 0.1, 0.2. As WSLS strategies are neither ZD nor unconditional strategies, the payoff relationships are not linear irrespective of errors and a discount factor.

Figures 3B and F show the case with an Extortioner strategy vs. $1000 + 2$ strategies. If there are no errors, Extortioner is unbeatable against any opponent as shown by black dots. For instance, when $w = 1$ and $\epsilon + \xi = 0$, Extortioner $\boldsymbol{p} = (0.86, 0.77, 0.09, 0)$ which passes over (P_E, P_E) can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ (black dots in Fig. 3B). Even if $w = 0.9$ and $\epsilon + \xi = 0$, Extortioner $\boldsymbol{p} = (0.955556, 0.855556, 0.1, 0; 0)$ which passes over (P_E, P_E) can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ (black dots in Fig. 3F).

However, as shown in Hao et. al [52] and Mamiya and Ichinose [53], when there are errors, there exists the region that the expected payoff of the Extortioner is lower than the opponent's payoff near (P_E, P_E) even though the increase of the Extortioner is still larger than the opponent due to $\chi > 1$ when the opponent tries to increase his payoff. Hao et. al called it *contingent extortion* [52]. When $w = 1$ and $\epsilon + \xi = 0.1$, Extortioner $\boldsymbol{p} = (0.926875, 0.818125, 0.111875, 003125)$ which passes over $(P_E + 0.1, P_E + 0.1)$ can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ (yellow-green dots in Fig. 3B). Even if $w = 0.9$ and $\epsilon + \xi = 0.1$, Extortioner $\boldsymbol{p} = (0.941667, 0.7, 0.241667, 0; 0)$ which has the same slope $\chi = 15$ can still exist

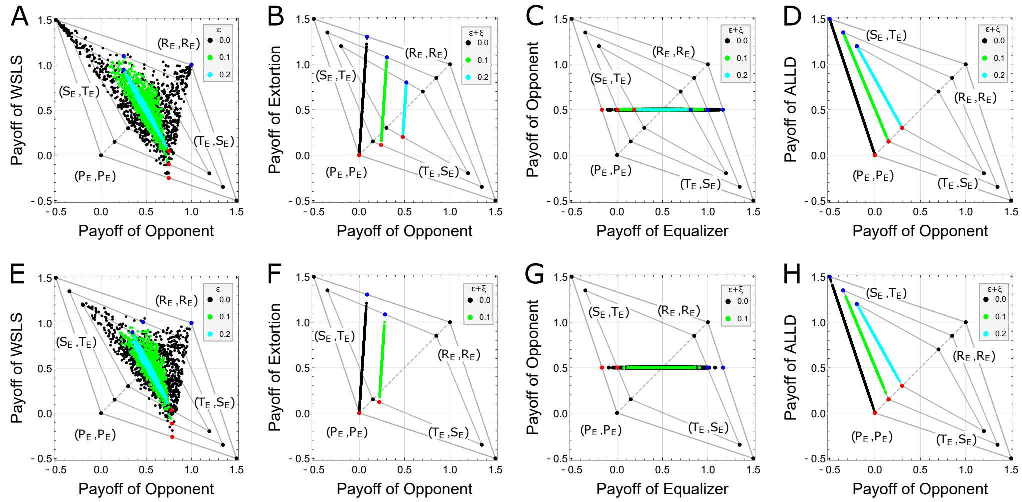

Fig 3. The payoff relationships between two players in the RPD game under observation errors. Payoff vector: $(T, R, P, S) = (1.5, 1, 0, -0.5)$. (A,E) WSLS strategy vs. $1000 + 2$ strategies. (B,F) Extortioner strategy vs. $1000 + 2$ strategies. (C,G) Equalizer strategy vs. $1000 + 2$ strategies. (D,H) ALLD strategy vs. $1000 + 2$ strategies. (A)–(D) are the case of $w = 1$ (no discount factor), and (E)–(H) correspond to (A)–(D) when $w = 0.9$, respectively.

(yellow-green dots in Fig. 3F). Nevertheless, these two Extortioners' expected payoffs are lower than the opponents near (P_E, P_E). When $w = 1$ and $\epsilon + \xi = 0.2$, Extortioner $\boldsymbol{p} = (1, 0.86, 0.14, 0)$ which passes over $(P_E + 0.2, P_E + 0.2)$ can impose a linear payoff relationship to the opponent, with the slope $\chi = 15$ (cyan dots in Fig. 3F). However, this Extortioner's payoff is lower than the opponent near (P_E, P_E), too. When $w = 0.9$ and $\epsilon + \xi = 0.2$, there is no Extortioner as shown in Fig. 2

Figures 3C and G show the case with an Equalizer strategy vs. $1000 + 2$ strategies. In those figures, we replaced the axes where the horizontal axis corresponds to Equalizer and the vertical axis corresponds to the opponent. When $w = 1$ and $\epsilon + \xi = 0, 0.1$ and 0.2, Equalizers $\boldsymbol{p} = (2/3, 1/3, 2/3, 1/3)$, $\boldsymbol{p} = (0.8, 0.365217, 0.634783, 0.2)$, and $\boldsymbol{p} = (0.99, 0.74, 0.26, 0.01)$ can fix the opponents' (player Y) expected payoffs at $s_Y = 0.5$ irrespective of Y's strategies, respectively (black, yellow-green, and cyan dots in Fig. 3C). Also, when $w = 0.9$ and $\epsilon + \xi = 0$ and 0.1, Equalizers $\boldsymbol{p} = (2/3, 0.277778, 0.722222, 1/3; 1/2)$, and $\boldsymbol{p} = (0.833333, 0.350242, 0.649758, 1/6; 1/2)$ can fix the opponents' (player Y) expected payoffs at $s_Y = 0.5$ irrespective of Y's strategies, respectively (black and yellow-green dots in Fig. 3G). When $w = 0.9$ and $\epsilon + \xi = 0.2$, there exists no Equalizer as shown in Fig. 1.

Lastly, we show the case of ALLD (Figs. 3D and H). ALLD strategy is one of the unconditional strategies where we set $r = 0$ in $\boldsymbol{p} = (r, r, r, r; r), 0 \leq r \leq 1$. As shown in Eq. (43) and Figs. 3D and H, w does not affect the expected payoff between both players. When $\epsilon + \xi = 0, 0.1$, and 0.2, those linear equations are $s_X + 3s_Y = 0$ (black), $s_X + 2.4s_Y - 0.51 = 0$ (yellow-green), and $s_X + 1.8s_Y - 0.84 = 0$ (cyan), respectively [53].

Conclusion

We considered both a discount factor and observation errors in the RPD game and analytically studied the strategies that enforce linear payoff relationships in the game.

First, we successfully derived the determinant form of the two players' expected payoffs even though a discount factor and observation errors are incorporated. Then, we searched for all possible strategies that enforce linear payoff relationships in the RPD game. As a result, we found both ZD strategies and unconditional strategies are the only strategy sets to enforce the relationship to the opponent. Moreover, we numerically showed that minimum discount rates for Equalizer ($\chi \to \infty$) and minimum extortion factors for other subsets of ZD strategies ($1 \leq \chi \leq \infty$) above which those ZD strategies exist.

Our results are limited to the two player RPD games. Other studies have focused on n-player games [19, 26, 27, 56]. It is worth investigating games including observation errors and a discount factor for n-player games . On the other hand, regarding memory, our study only used memory-1 strategies. A recent study revealed the role of longer memories for the evolution of cooperation, which is another direction to investigate [51].

When spatial structures are included, the different role of the Extortioner is known [16–18, 21]. Extortioners are tied with ALLDs. Thus, Extortioners can neutrally invade the sea of ALLDs in a spatial structure. On the other hand, the best response to Extortioners is ALLC. Once ALLC happens, the clusters of ALLC are better than those of the Extortioner. Then, cooperation is promoted. In this way, it has been demonstrated that Extortion acts as a catalyst for cooperation. Another interest is how observation errors and a discount factor affect the evolution of cooperation in a spatial setting. Our study contributes to open various new research directions of ZD strategies.

Supporting information

S1 Appendix. Proof of $v_1 + v_2 + v_3 + v_4 = 1$.

We show the sum of elements in the mean distribution $\boldsymbol{v} = (v_1, v_2, v_3, v_4)$ is equal to one. We define

$$\boldsymbol{v} = (1-w)\boldsymbol{v}(0)\sum_{t=0}^{\infty}(wM)^t. \tag{33}$$

This is another form of Eq (13). Because the sum of every row in the transition matrix M is equal to one, the sum of every row of $\sum_{t=0}^{\infty}(wM)^t$ is equal to $1/(1-w)$. The sum of vector elements in $\boldsymbol{v}(0)$ is unchanged from 1 even if the vector is multiplied by $(1-w)\sum_{t=0}^{\infty}(wM)^t$. Therefore, $v_1 + v_2 + v_3 + v_4 = 1$ holds.

S2 Appendix. Calculation of $\boldsymbol{v}(0) = \boldsymbol{v}^T M_0$.

We show that $\boldsymbol{v}(0)$ and $\boldsymbol{v}^T M_0$ are equal. $\boldsymbol{v}$ and M_0 are defined by Eq. (13) and Eq. (14), respectively. We calculate the matrix multiplication $\boldsymbol{v}^T M_0$.

$$\begin{aligned}
\boldsymbol{v}^T M_0 &= (v_1, v_2, v_3, v_4)\begin{pmatrix} p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \\ p_0q_0 & p_0(1-q_0) & (1-p_0)q_0 & (1-p_0)(1-q_0) \end{pmatrix} \\
&= (p_0q_0(v_1+v_2+v_3+v_4), p_0(1-q_0)(v_1+v_2+v_3+v_4), \\
&\qquad (1-p_0)q_0(v_1+v_2+v_3+v_4), (1-p_0)(1-q_0)(v_1+v_2+v_3+v_4)) \\
&= (p_0q_0, p_0(1-q_0), (1-p_0)q_0, (1-p_0)(1-q_0)) = \boldsymbol{v}(0)
\end{aligned} \tag{34}$$

Therefore, the following holds:

$$\boldsymbol{v}(0) = \boldsymbol{v}^T M_0. \tag{35}$$

S3 Appendix. Strategies that enforce $D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \boldsymbol{1}) = 0$.

To search for all possible strategies that make $D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \boldsymbol{1}) = 0$, we express Eq. (25) in component form:

$$s\begin{pmatrix} w(\tau p_1 q_1 + \epsilon p_1 q_2 + \epsilon p_2 q_1 + \xi p_2 q_2) - 1 + p_0 q_0 (1-w) \\ w(\epsilon p_1 q_3 + \xi p_1 q_4 + \tau p_2 q_3 + \epsilon p_2 q_4) + p_0 q_0 (1-w) \\ w(\epsilon p_3 q_1 + \tau p_3 q_2 + \xi p_4 q_1 + \epsilon p_4 q_2) + p_0 q_0 (1-w) \\ w(\xi p_3 q_3 + \epsilon p_3 q_4 + \epsilon p_4 q_3 + \tau p_4 q_4) + p_0 q_0 (1-w) \end{pmatrix} + t\begin{pmatrix} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) \\ w(\mu p_3 + \eta p_4) + p_0(1-w) \\ w(\eta p_3 + \mu p_4) + p_0(1-w) \end{pmatrix}$$
$$+u\begin{pmatrix} w(\mu q_1 + \eta q_2) - 1 + q_0(1-w) \\ w(\mu q_3 + \eta q_4) + q_0(1-w) \\ w(\eta q_1 + \mu q_2) - 1 + q_0(1-w) \\ w(\eta q_3 + \mu q_4) + q_0(1-w) \end{pmatrix} + v\begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \boldsymbol{0} \tag{36}$$

By taking out $\boldsymbol{q}$ from Eq. (36), we obtain

$$\begin{pmatrix} w((s(\tau p_1 + \epsilon p_2) + u\mu) q_1 + (s(\epsilon p_1 + \xi p_2) + u\eta) q_2) + (s p_0 + u)(1-w) q_0 \\ w((s(\epsilon p_1 + \tau p_2) + u\mu) q_3 + (s(\xi p_1 + \epsilon p_2) + u\eta) q_4) + (s p_0 + u)(1-w) q_0 \\ w((s(\epsilon p_3 + \xi p_4) + u\eta) q_1 + (s(\tau p_3 + \epsilon p_4) + u\mu) q_2) + (s p_0 + u)(1-w) q_0 \\ w((s(\xi p_3 + \epsilon p_4) + u\eta) q_3 + (s(\epsilon p_3 + \tau p_4) + u\mu) q_4) + (s p_0 + u)(1-w) q_0 \end{pmatrix}$$
$$+t\begin{pmatrix} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) \\ w(\mu p_3 + \eta p_4) + p_0(1-w) \\ w(\eta p_3 + \mu p_4) + p_0(1-w) \end{pmatrix} + \begin{pmatrix} -s-u \\ 0 \\ -u \\ 0 \end{pmatrix} + v\begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \boldsymbol{0}. \tag{37}$$

Here, we search for strategies which satisfy $D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \boldsymbol{1}) = 0$ irrespective of Y's strategy $\boldsymbol{q}$, meaning that Eq. (37) must hold true irrespective of $\boldsymbol{q}$. Therefore, the coefficients of each element $\boldsymbol{q}$ in Eq. (37) must equal zero, that is, the following conditions are necessary:

$$\begin{cases} w(s(\epsilon p_1 + \xi p_2) + u\eta) & = 0 \\ w(s(\epsilon p_3 + \xi p_4) + u\eta) & = 0 \\ w(s(\tau p_1 + \epsilon p_2) + u\mu) & = 0 \\ w(s(\tau p_3 + \epsilon p_4) + u\mu) & = 0 \\ w(s(\epsilon p_1 + \tau p_2) + u\mu) & = 0 \\ w(s(\xi p_1 + \epsilon p_2) + u\eta) & = 0 \\ w(s(\xi p_3 + \epsilon p_4) + u\eta) & = 0 \\ w(s(\epsilon p_3 + \tau p_4) + u\mu) & = 0 \\ (s p_0 + u)(1-w) & = 0. \end{cases} \tag{38}$$

When Eq. (38) holds, the first terms of Eq. (37) are eliminated and we obtain

$$t\begin{pmatrix} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) \\ w(\mu p_3 + \eta p_4) + p_0(1-w) \\ w(\eta p_3 + \mu p_4) + p_0(1-w) \end{pmatrix} + \begin{pmatrix} -s-u \\ 0 \\ -u \\ 0 \end{pmatrix} + v\begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \boldsymbol{0}. \tag{39}$$

If there exist real numbers, $s, t, u, v, \alpha, \beta$, and γ such that Eq. (38) and Eq. (39) are satisfied simultaneously, $D(\boldsymbol{p}, \boldsymbol{q}, \alpha \boldsymbol{S}_X + \beta \boldsymbol{S}_Y + \gamma \boldsymbol{1}) = 0$ holds irrespective of $\boldsymbol{q}$. We first solve Eq. (38). After some calculations, Eq. (38) becomes

$$\begin{cases} ws(\epsilon-\xi)(p_1-p_2) &= 0\\ ws(\epsilon-\xi)(p_3-p_4) &= 0\\ ws(1-3\epsilon-\xi)(p_1-p_2) &= 0\\ ws(1-3\epsilon-\xi)(p_3-p_4) &= 0\\ w(s(\epsilon p_1+\tau p_2)+u\mu) &= 0\\ w(s(\xi p_1+\epsilon p_2)+u\eta) &= 0\\ w(s(\xi p_3+\epsilon p_4)+u\eta) &= 0\\ w(s(\epsilon p_3+\tau p_4)+u\mu) &= 0\\ (sp_0+u)(1-w) &= 0. \end{cases} \tag{40}$$

When we solve the first four equations, we obtain (1)$w=0$, (2) $s=0$, (3) $\epsilon-\xi=0$ and $1-3\epsilon-\xi=0$, (4) $p_1-p_2=0$ and $p_3-p_4=0$. We further analyze whether these solutions satisfy the last four equations and Eq. (39) by dividing them into four cases as follows.

Case (1) $w=0$:

In this case, we substitute $w=0$ into Eq. (40) to obtain

$$sp_0+u=0. \tag{41}$$

Therefore one of the solutions of Eq. (40) is $w=0$ and $u=-sp_0$. Next, we check whether this solution satisfies Eq. (39). We substitute $w=0$ and $u=-sp_0$ into Eq. (39) to obtain

$$s\begin{pmatrix} p_0-1\\ 0\\ p_0\\ 0\end{pmatrix}+t\begin{pmatrix} p_0-1\\ p_0-1\\ p_0\\ p_0\end{pmatrix}+v\begin{pmatrix}\alpha R_E+\beta R_E+\gamma\\ \alpha S_E+\beta T_E+\gamma\\ \alpha T_E+\beta S_E+\gamma\\ \alpha P_E+\beta P_E+\gamma\end{pmatrix}=\mathbf{0}. \tag{42}$$

There exist real numbers s,t,u,v,α,β, and γ which satisfies Eq. (42) as follows:

$$\begin{aligned}
s &= \frac{v\alpha(S_E(-P_E-R_E+S_E)+T_E(P_E+R_E-T_E))}{(1-p_0)(P_E-S_E)+p_0(T_E-R_E)}\\
t &= \frac{v\alpha(S_E(2P_E-S_E+p_0(-P_E-R_E+S_E))+T_E(-2P_E+T_E+p_0(P_E+R_E-T_E)))}{(1-p_0)(P_E-S_E)+p_0(T_E-R_E)}\\
u &= -sp_0\\
\beta &= \frac{\alpha((1-p_0)(T_E-P_E)+p_0(R_E-S_E))}{(1-p_0)(P_E-S_E)+p_0(T_E-R_E)}\\
\gamma &= \frac{\alpha(S_E-T_E)((-1+p_0)^2P_E+p_0(1-p_0)(T_E+S_E)+p_0^2R_E)}{(1-p_0)(P_E-S_E)+p_0(T_E-R_E)}\\
&\forall v,\alpha.
\end{aligned} \tag{43}$$

Thus, when $w=0$ (one-shot game), two player's payoffs always become linear irrespective of $\boldsymbol{p}_0$.

Case (2) $s = 0$:

In this case, we substitute $s = 0$ into Eq. (40) to obtain

$$\begin{cases} u\mu & = 0 \\ u\eta & = 0 \\ u(1-w) & = 0. \end{cases} \tag{44}$$

The equations $\mu = 0$ and $\eta = 0$ do not hold at the same time due to $\mu = 1 - \epsilon - \xi$ and $\eta = \epsilon + \xi$. Therefore, one of the solutions of Eq. (40) is $s = 0$ and $u = 0$. Next, we check whether this solution satisfies Eq. (39). We substitute $s = 0$ and $u = 0$ into Eq. (39) to obtain

$$t \begin{pmatrix} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) \\ w(\mu p_3 + \eta p_4) + p_0(1-w) \\ w(\eta p_3 + \mu p_4) + p_0(1-w) \end{pmatrix} + v \begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \mathbf{0}. \tag{45}$$

Here, when we set $t = 0$, either equation

$$v = 0 \tag{46}$$

or

$$\begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \mathbf{0} \tag{47}$$

must hold. When we set $v = 0$, we obtain the trivial solution $(s, t, u, v) = (0, 0, 0, 0)$. Also, when we solve Eq. (47), we obtain the trivial solution $(\alpha, \beta, \gamma) = (0, 0, 0)$. Hence, we do not have to consider the case of $t = 0$. Therefore, in the following, we only consider $t \neq 0$. Replacing constants $-\alpha v/t$, $-\beta v/t$, and $-\gamma v/t$ with α, β, and γ, we obtain,

$$\begin{aligned} w(\mu p_1 + \eta p_2) - 1 + p_0(1-w) &= \alpha R_E + \beta R_E + \gamma \\ w(\eta p_1 + \mu p_2) - 1 + p_0(1-w) &= \alpha S_E + \beta T_E + \gamma \\ w(\mu p_3 + \eta p_4) + p_0(1-w) &= \alpha T_E + \beta S_E + \gamma \\ w(\eta p_3 + \mu p_4) + p_0(1-w) &= \alpha P_E + \beta P_E + \gamma. \end{aligned} \tag{48}$$

If there exist α, β, and γ satisfying Eq. (48), there must be solutions that Eq. (25) holds. This solution corresponds to ZD strategies with observation errors and a discount factor. This is consistent with Eq. (6) in [52] when $w = 1$.

Case (3) $\epsilon - \xi = 0$ and $1 - 3\epsilon - \xi = 0$:

In this case, the equations $\epsilon - \xi = 0$ and $1 - 3\epsilon - \xi = 0$ lead to $\epsilon = 1/4$ and $\xi = 1/4$. When $\epsilon = 1/4$ and $\xi = 1/4$, the expected payoffs $R_E = 1/2(R+S), S_E = 1/2(R+S), T_E = 1/2(T+P)$, and $P_E = 1/2(T+P)$ hold, which do not satisfy the condition of the prisoner's dilemma game: $T_E > R_E > P_E > S_E$. Hence, we can exclude this solution.

Case (4) $p_1 - p_2 = 0$ and $p_3 - p_4 = 0$:

In this case, we substitute $p_1 - p_2 = 0$ and $p_3 - p_4 = 0$ into Eq. (40) to obtain

$$\begin{cases} w\mu(sp_1 + u) & = 0 \\ w\eta(sp_1 + u) & = 0 \\ w\eta(sp_3 + u) & = 0 \\ w\mu(sp_3 + u) & = 0 \\ (sp_0 + u)(1 - w) & = 0. \end{cases} \tag{49}$$

Because the equations $\mu = 0$ and $\eta = 0$ do not hold at the same time and $w \neq 0$, we obtain

$$\begin{cases} sp_1 + u & = 0 \\ sp_3 + u & = 0 \\ (sp_0 + u)(1 - w) & = 0. \end{cases} \tag{50}$$

Therefore, we obtain two solutions $p_0 = p_1 = p_2 = p_3 = p_4 = -u/s$ or $p_1 = p_2 = p_3 = p_4 = -u/s$ and $w = 1$. Both solutions are called unconditional strategies [14, 55]. The former represents unconditional strategies in the case of $w \neq 1$. The latter represents unconditional strategies in the case of $w = 1$. Next, we check whether this solution satisfies Eq. (39). We substitute the former solution $p_0 = p_1 = p_2 = p_3 = p_4$ and $u = -sp_0$ into Eq. (39) to obtain

$$s \begin{pmatrix} p_0 - 1 \\ 0 \\ p_0 \\ 0 \end{pmatrix} + t \begin{pmatrix} w(\mu p_0 + \eta p_0) - 1 + p_0(1 - w) \\ w(\eta p_0 + \mu p_0) - 1 + p_0(1 - w) \\ w(\mu p_0 + \eta p_0) + p_0(1 - w) \\ w(\eta p_0 + \mu p_0) + p_0(1 - w) \end{pmatrix} + v \begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \mathbf{0}. \tag{51}$$

According to $\mu + \eta = 1$, we obtain

$$s \begin{pmatrix} p_0 - 1 \\ 0 \\ p_0 \\ 0 \end{pmatrix} + t \begin{pmatrix} p_0 - 1 \\ p_0 - 1 \\ p_0 \\ p_0 \end{pmatrix} + v \begin{pmatrix} \alpha R_E + \beta R_E + \gamma \\ \alpha S_E + \beta T_E + \gamma \\ \alpha T_E + \beta S_E + \gamma \\ \alpha P_E + \beta P_E + \gamma \end{pmatrix} = \mathbf{0}. \tag{52}$$

Eq. (52) corresponds to Eq. (42). Therefore, there exist real numbers $s, t, u, v, \alpha, \beta$, and γ which satisfies Eq. (52) as Eq. (43). Finally, we substitute the latter solution $p_1 = p_2 = p_3 = p_4 = -u/s$ and $w = 1$ into Eq (39), we obtain the same real numbers $s, t, u, v, \alpha, \beta$, and γ in the case of the former solution.

This strategy set corresponds to unconditional strategies $\boldsymbol{p} = (r, r, r, r; r), 0 \leq r \leq 1$. Therefore, the unconditional strategies enforce a linear payoff relationship in the RPD game with both observation errors and a discount factor because there exist real numbers $s, t, u, v, \alpha, \beta$, and γ such that Eq. (39) and Eq. (40) are satisfied.

Acknowledgments

This study was partly supported by the HAYAO NAKAYAMA Foundation for Science & Technology and Culture and JSPS KAKENHI Grant Number JP19K04903 (G.I.).

References

1. Mailath J, Samuelson L. Repeated Games and Reputation. Oxford: Oxford University Press; 2006.
2. Trivers RL. The evolution of reciprocal altruism. Q Rev Biol. 1971;46:35–57.
3. Nowak MA. Evolutionary Dynamics. Cambridge, MA: Harvard University Press; 2006.
4. Sigmund K. The Calculus of Selfishness. Princeton, NJ: Princeton University Press; 2010.
5. Smith JM. Evolution and the Theory of Games. Cambridge: Cambridge University Press; 1982.
6. Nowak MA, Sigmund K. Tit for tat in heterogeneous populations. Nature. 1992;355:250–253.
7. Nowak M, Sigmund K. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature. 1993;364:56–58.
8. Kraines D, Kraines V. Learning to cooperate with Pavlov an adaptive strategy for the iterated Prisoner's Dilemma with noise. Theory Decis. 1993;35:107–150.
9. Press WH, Dyson FJ. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. Proc Natl Acad Sci USA. 2012;109:10409–10413.
10. Stewart AJ, Plotkin JB. Extortion and cooperation in the Prisoner's Dilemma. Proc Natl Acad Sci USA. 2012;109:10134–10135.
11. Akin E. The iterated prisoner's dilemma: Good strategies and their dynamics. Ergodic Theory, Advances in Dynamics. 2016;(de Gruyter, Berlin, 2016):77–107.
12. Adami C, Hintze A. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. Nat Comm. 2013;4:2193.
13. Hilbe C, Nowak MA, Sigmund K. Evolution of extortion in Iterated Prisoner's Dilemma games. Proc Natl Acad Sci USA. 2013;110:6913–6918.
14. Hilbe C, Nowak MA, Traulsen A. Adaptive dynamics of extortion and compliance. PLoS ONE. 2013;8:e77886.
15. Chen J, Zinger A. The robustness of zero-determinant strategies in Iterated Prisoner's Dilemma games. J Theor Biol. 2014;357:46–54.
16. Szolnoki A, Perc M. Evolution of extortion in structured populations. Phys Rev E. 2014;89:022804.
17. Szolnoki A, Perc M. Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. Sci Rep. 2014;4:5496.
18. Wu ZX, Rong Z. Boosting cooperation by involving extortion in spatial prisoner's dilemma games. Phys Rev E. 2014;90:062102.
19. Hilbe C, Wu B, Traulsen A, Nowak MA. Evolutionary performance of zero-determinant strategies in multiplayer games. J Theor Biol. 2015;374:115–124.

20. Liu J, Li Y, Xu C, Hui PM. Evolutionary behavior of generalized zero-determinant strategies in iterated prisoner's dilemma. Physica A. 2015;430:81–92.

21. Xu X, Rong Z, Wu ZX, Zhou T, Tse CK. Extortion provides alternative routes to the evolution of cooperation in structured populations. Phys Rev E. 2017;95:052302.

22. Wang J, Guo J. A synergy of punishment and extortion in cooperation dilemmas driven by the leader. Chaos, Solitons & Fractals. 2019;119:263–268.

23. Stewart AJ, Plotkin JB. From extortion to generosity, evolution in the Iterated Prisoner's Dilemma. Proc Natl Acad Sci USA. 2013;110:15348–15353.

24. Mao Y, Xu X, Rong Z, Wu ZX. The emergence of cooperation-extortion alliance on scale-free networks with normalized payoff. EPL. 2018;122:50005.

25. Xu X, Rong Z, Tian Z, Wu ZX. Timescale diversity facilitates the emergence of cooperation-extortion alliances in networked systems. Neurocomputing. 2019;350:195–201.

26. Hilbe C, Wu B, Traulsen A, Nowak MA. Cooperation and control in multiplayer social dilemmas. Proc Natl Acad Sci USA. 2014;111:16425–16430.

27. Pan L, Hao D, Rong Z, Zhou T. Zero-determinant strategies in iterated public goods game. Sci Rep. 2015;5:13096.

28. Milinski M, Hilbe C, Semmann D, Sommerfeld R, Marotzke J. Humans choose representatives who enforce cooperation in social dilemmas through extortion. Nat Comm. 2016;7:10915.

29. Stewart AJ, Parsons TL, Plotkin JB. Evolutionary consequences of behavioral diversity. Proc Natl Acad Sci USA. 2016;113:E7003–E7009.

30. McAvoy A, Hauert C. Autocratic strategies for iterated games with arbitrary action spaces. Proc Natl Acad Sci USA. 2016;113:3573–3578.

31. McAvoy A, Hauert C. Autocratic strategies for alternating games. Theor Popul Biol. 2017;113:13–22.

32. Taha MA, Ghoneim A. Zero-determinant strategies in repeated asymmetric games. Appl Math Comput. 2019; p. 124862.

33. Engel A, Feigel A. Single equalizer strategy with no information transfer for conflict escalation. Phys Rev E. 2018;98:012415.

34. Hilbe C, Röhl T, Milinski M. Extortion subdues human players but is finally punished in the prisoner's dilemma. Nat Comm. 2014;5:3976.

35. Wang Z, Zhou Y, Lien JW, Zheng J, Xu B. Extortion can outperform generosity in the iterated prisoner's dilemma. Nat Comm. 2016;7:11125.

36. Hilbe C, Hagel K, Milinski M. Asymmetric power boosts extortion in an economic experiment. PLoS ONE. 2016;11:e0163867.

37. Becks L, Milinski M. Extortion strategies resist disciplining when higher competitiveness is rewarded with extra gain. Nat Comm. 2019;10:783.

38. Akin E. What you gotta know to play good in the Iterated Prisoner's Dilemma. Games. 2015;6:175–190.

39. Hilbe C, Chatterjee K, Nowak MA. Partners and rivals in direct reciprocity. Nat Hum Behav. 2018;2:469–477.

40. Hu Q, Wang S, Cheng X, Ma L, Bie R. Solving the Crowdsourcing Dilemma Using the Zero-Determinant Strategies. IEEE Trans Inf Forensic Secur. 2019;Early Access:1.

41. Hu Q, Wang S, Ma P, Cheng X, Lv W, Bie R. Quality Control in Crowdsourcing Using Sequential Zero-Determinant Strategies. IEEE Trans Knowl Data Eng. 2019;Early Access:1.

42. Tang C, Li X, Cao M, Zhang Z, Yu X. Incentive Mechanism for Macrotasking Crowdsourcing: A Zero-Determinant Strategy Approach. IEEE Internet Things J. 2019;6:8589–8601.

43. Kandori M. Introduction to repeated games with private monitoring. J Econ Theor. 2002;102:1–15.

44. Nowak MA, Sigmund K, El-Sedy E. Automata, repeated games and noise. J Math Biol. 1995;33:703–722.

45. Fudenberg D, Rand DG, Dreber A. Slow to anger and fast to forgive: Cooperation in an uncertain world. Am Econ Rev. 2012;102:720–749.

46. Fudenberg D, Levine D, Maskin E. The folk theorem with imperfect public information. Econometrica. 1994;62:997–1039.

47. Sekiguchi T. Efficiency in repeated prisoner's dilemma with private monitoring. J Econ Theor. 1997;76:345–361.

48. Barlo M, Carmona G, Sabourian H. Repeated games with one-memory. J Econ Theor. 2009;144:312–336.

49. Mailath GJ, Morris S. Repeated games with almost-public monitoring. J Econ Theor. 2002;102:189–228.

50. Mailath GJ, Olszewski W. Folk theorems with bounded recall under (almost) perfect monitoring. Gam Econ Behav. 2011;71:174–192.

51. Hilbe C, Martinez-Vaqueroand LA, Chatterjee K, Nowak MA. Memory-n strategies of direct reciprocity. Proc Natl Acad Sci USA. 2017;114:4715–4720.

52. Hao D, Rong Z, Zhou T. Extortion under uncertainty: Zero-determinant strategies in noisy games. Phys Rev E. 2015;91:052803.

53. Mamiya A, Ichinose G. Strategies that enforce linear payoff relationships under observation errors in Repeated Prisoner's Dilemma game. J Theor Biol. 2019;477:63–76.

54. Hilbe C, Traulsen A, Sigmund K. Partners or rivals? Strategies for the iterated prisoner's dilemma. Games Econ Behav. 2015;92:41–52.

55. Ichinose G, Masuda N. Zero-determinant strategies in finitely repeated games. J Theor Biol. 2018;438:61–77.

56. Govaert A, Cao M. Zero-Determinant strategies in finitely repeated n-player games. arXiv. 2019;.