Joosten, Reinoud

Working Paper

A small Fish War: an example with frequency-dependent stage payoffs

Papers on Economics and Evolution, No. 0506

Provided in Cooperation with:
Max Planck Institute of Economics

PAPERS on Economics & Evolution

MAX-PLANCK-GESELLSCHAFT

0506

A small Fish War: An example with frequency-dependent stage payoffs

by

Reinoud Joosten

The *Papers on Economics and Evolution* are edited by the Evolutionary Economics Group, MPI Jena. For editorial correspondence, please contact: evopapers@mpiew-jena.mpg.de

ISSN 1430-4716

Max Planck Institute for
Research into Economic Systems
Evolutionary Economics Group
Kahlaische Str. 10
07745 Jena, Germany
Fax: ++49-3641-686868

A small Fish War: an example with frequency-dependent stage payoffs*

Reinoud Joosten†

May 20, 2005

Abstract

Two agents possess the fishing rights to a lake. Each period they have two options, to catch without restraint, e.g., to use a fine-mazed net, or to catch with some restraint, e.g., to use a wide-mazed net. The use of a fine-mazed net always yields a higher immediate catch than the alternative. The present catches depend on the behavior of the agents in the past. The more often the agents have used the fine-mazed net in the past, the lower the present catches are independent from the type of nets being used.

We determine feasible rewards and provide (subgame perfect) equilibria for the limiting average reward criterion using methods inspired by the repeated-games literature. Our analysis shows that a 'tragedy of the commons' can be averted, as sustainable Pareto-efficient outcomes can be supported by subgame perfect equilibria.
JEL-codes C72, C73, Q20, Q22
Keywords games with frequency-dependent stage payoffs, limiting average reward, equilibria, renewable common-pool resources

1 Introduction

There exists a vast literature on fishery games following the seminal paper *The Great Fish War* by Levhari & Mirman [1980], where a non-cooperative difference game was used to model strategic interaction between two agents exploiting a natural resource. The paper revealed that under various regimes of strategic interaction, e.g., Cournot-Nash, Stackelberg or collusion, the agents over-exploit the natural resource, albeit in different degrees. So, strategic behavior in a fishery game may very well lead to effects which bear resemblance to the famous 'tragedy of the commons' (Hardin [1968]).

*Part of this paper's research was performed during a stay at the Max Planck Institute for research into Economic Systems in Jena. The author thanks George von Wangenheim and Ulrich Witt for their hospitality, and Werner Güth for criticism to related work.

†FELab & School of Business, Management and Technology, University of Twente, POB 217, 7500 AE Enschede, The Netherlands. Email: r.a.m.g.joosten@utwente.nl

The Great Fish War was presented as a discrete-time analog of a differential game. A large part of the ensuing literature also took a difference/differential game approach and the methods of analysis were, not surprisingly, found in the corresponding tool boxes. However, Rabah Amir showed that this game can also be interpreted as a stochastic game with uncountable state and action spaces (Amir [2003]). Shapley [1953] introduced a stochastic game in which several states are distinguished and in which at each stage of the play the players choose an action in the game of the current state. The choices made by the players determine their payoffs in the stage game, as well as a probability distribution over the states as to where the play will continue. For overviews and results on differential games and stochastic games, we refer to e.g., Dockner *et al.* [2001] and to e.g., Vrieze [1987], Thuijsman [1992], and Vieille [2000a,b].

Fisheries fall into the category of renewable common-pool resources[1] to which difference/differential games can be applied to model strategic interaction as well (see e.g., Hanley *et al.* [1997]). Alternative game-theoretical approaches to model fisheries can also be found in industrial organization (see e.g., Shy [1995]), and the intersection of economics and the political sciences (see e.g., Ostrom *et al.* [1994]). In the latter two approaches however, the modeling of inter-temporal aspects is usually not as explicit as in the ones using differential or stochastic games.

Common-pool resource systems are often associated with social dilemmas, or social traps (see e.g., Dawes [1975,1980]), Platt [1973]). Dawes [1980] states that two properties define a social dilemma: (a) each individual receives a higher payoff for a socially defecting choice than for a socially cooperative choice, no matter what other individuals in society do, and (b) all individuals are better off if all cooperate than if all defect. A social trap is a social dilemma, but with an effect in the time-dimension. Certain behavior leads to a small positive outcome which is immediate, and a large negative outcome which is delayed (see Platt [1973]). For an excellent overview on social dilemmas and related issues see e.g., Komorita & Parks [1994].

The aim of this paper is to introduce and provide a first analysis of a new type of fishery game. Therefore, we make some concessions to reality. For instance, we do not incorporate prices, which in essence is an assumption that prices remain constant, or alternatively that the entire catch is consumed directly; we do not model strategic interaction on the markets, which in essence boils down to an assumption that the market is not oligopolistic on the supply side; finally, in modeling the environment we abstract from stochasticity. Instead, we have the following stylized setup.

Two agents possess the fishing rights to a lake. Each period each player has essentially two options, to fish with no restraint at all, e.g., use a fine-mazed net, or to fish with some restraint, e.g., use a wide-mazed net. As may

[1] See Ostrom [2000] on the terminology 'common-property resource' (Gordon [1954]).

be expected the use of a fine-mazed net always yields a higher immediate catch. However, the immediate catches also depend on the behavior of the agents in the past. The more frequently the agents have used the fine-mazed net in the past, the lower the immediate catches are independent from the type of nets currently in use. The reason for this phenomenon is that in using the fine-mazed net, also young fish are caught, which do not grow up to adulthood. Hence, in future catches these fish are missing, moreover, these fish can not reproduce. We assume that each agent wishes to maximize the average size of the catch over an infinite time-horizon.

We model a Fish War as a game with frequency-dependent stage payoffs, or FD-game (Joosten *et al.* [2003]). The underlying idea stems from the work of economist and psychologist Herrnstein on distributed choice in which stimuli changed depending on the choices made by experimental subjects (cf., e.g., Herrnstein [1997]). Until recently, these effects had been examined primarily in 'games against nature' settings. The first contribution using the concept of frequency-dependent payoffs in a multi-person strategic-interaction framework was Brenner & Witt [2003].

In Joosten *et al.* [2003] a formal definition of FD-games was given and an extensive analysis was undertaken. FD-games are stochastic games with finite action spaces and infinite state spaces, but the analysis of infinitely repeated games can very well be adapted and generalized to this type of games. Several Folk Theorems were derived which imply that if the agents are sufficiently patient, all individually-rational rewards can be supported by an equilibrium involving threats.

This paper presents a rather significant innovation relative to the original FD-game framework, where the relative frequencies of past play translate into linear and additive quantitative effects on the stage payoffs. Here, we model the effects of the relative frequencies of past play on the stage payoffs in a non-linear and multiplicative mathematical relationship in order to obtain, in our view at least, a more appropriate scenario. The analysis itself and the results obtained are closely related to Joosten *et al.* [2003].

Our analysis shows that the 'tragedy of commons' does not seem inevitable, as Pareto-efficient outcomes can be sustained by subgame perfect equilibria.[2] In a setting where the stage payoffs deteriorate to a mere ten percent of the maximum level due to fishing without restraint, we find that in each Pareto-efficient equilibrium the catches remain significantly above minimal level. Both agents must catch with restraint for approximately 85.6% of the stages, in the remaining stages precisely one agent catches with restraint. In the symmetric Pareto-efficient equilibrium both receive slightly more than in the 'perfect restraint' equilibrium, but over seven times the symmetric minimal 'no restraint' rewards.

[2] In differential games history-dependent strategies may serve to avert a 'tragedy of the commons' as well (e.g., Tolwinski *et al.* [1986]).

FD-games offer attractive possibilities to model strategic interaction with payoffs changing in time, without having to venture deeply into the fields of differential or stochastic games. As Rabah Amir [2003] showed, connections can be shown to exist between stochastic games and difference games. However, the formulation of the models and especially the tool boxes for analysis for both types of games differ so significantly that it seems (still) justified to speak about separate and independent subfields of game theory. As such, FD-games, while belonging to the class of stochastic games, resemble in formulation of the model and the tool box for analysis the class of repeated games more than any other class of games.

Next, we introduce the new type of fishery game, in Section 3 we compute rewards of the game using the limiting average reward criterion. Section 4 deals with establishing rewards that can be associated with an equilibrium or a subgame perfect equilibrium. Section 5 concludes with a discussion. Proofs are relegated to the appendix.

2 The FD-Fish War

The Fish War with frequency-dependent stage payoffs, or FD-Fish War, is played by players A and B at discrete moments in time called stages. Each player has two actions and each stage each players independently and simultaneously chooses an action. We denote the action set of player A (B) by J^A (J^B) and $J \equiv J^A \times J^B$. Action 1 for either player will always denote the action without or with very little restraint, e.g., catching with fine-mazed net or catching a high quantity. The other action will always denote the action where there exists some upper limit or restriction, i.e., catching with wide-mazed nets or catching a low quantity. The payoffs at stage $t' \in \mathbb{N}$ of the play depend on the choices of the players at that stage, and on the relative frequencies with which all actions where actually chosen until then.

Let $h^A_{t'} = \left(j^A_1, ..., j^A_{t'-1}\right)$ be the sequence of actions chosen by player A until stage $t' \geq 2$ and let $q \geq 0$, then define ρ^A_t recursively for $t \leq t'$ by

$$\rho^A_1 = \rho^A \in [0,1]\,, \text{ and } \rho^A_t = \begin{cases} \frac{q+t-1}{q+t}\rho^A_{t-1} + \frac{1}{q+t} & \text{if } j^A_{t-1} = 1, \\ \frac{q+t-1}{q+t}\rho^A_{t-1} & \text{if } j^A_{t-1} = 2. \end{cases} \qquad (1)$$

Define ρ^B_t for player B similarly. Taking $q \gg 0$ serves to moderate 'early' effects on the stage payoffs. Recall that j^A_{t-1} denotes the action chosen by A at stage $t-1$, hence, the number ρ^A_t converges in the long run to the relative frequency with which A chose the first action before stage t, regardless of the numbers ρ^A and q.

At stage $t \in \mathbb{N}$, the players have chosen action sequences h^A_t and h^B_t which induce the numbers ρ^A_t and ρ^B_t. The latter two numbers determine the state in which the play is at stage t. Slightly more formal, we say that the play at stage $t \in \mathbb{N}$ is in state $s_t \equiv \left(\rho_t^A, \rho_t^B\right)$. Observe that Eq. (1) implies that there are four possible successor states s_{t+1} to state s_t depending on the action pair chosen by the players at stage t.

At each stage a bi-matrix game is played, and the choices of the players realized at that stage determine the stage payoffs to the players. Let

$$A = B^\top = \begin{bmatrix} \frac{11}{2} & 6 \\ \frac{7}{2} & 4 \end{bmatrix}.$$

Then, for given $\mu_t \in [0,1]$ at stage $t \in \mathbb{N}$, the stage payoffs are given by

$$\left(A\left(\mu_t\right), B\left(\mu_t\right)\right) = \mu_t(A,B) = \begin{bmatrix} \frac{11}{2}\mu_t, \frac{11}{2}\mu_t & 6\mu_t, \frac{7}{2}\mu_t \\ \frac{7}{2}\mu_t, 6\mu_t & 4\mu_t, 4\mu_t \end{bmatrix}.$$

With respect to the payoff matrices $\left(A\left(\mu_t\right), B\left(\mu_t\right)\right)$, we have the following rather standard interpretation. Given μ_t, suppose action pair $(i,j) \in J$ is chosen, then A receives a stage payoff equal to $a_{ij}\mu_t$ and B receives $b_{ij}\mu_t$. For instance, if player A uses action 2 and player B uses action 1, then A's stage payoff is $\frac{7}{2}\mu_t$ and B's stage payoff equals $6\mu_t$.

Observe that the first action dominates the second action strictly for all $\mu_t > 0$. This means that fishing without restraint almost always yields a strictly higher catch in the current stage than fishing with restraint. So, the unique Nash equilibrium in the stage game is the strategy pair in which both players use their first action. Note furthermore, that the setup of the stage game is such that negative externalities exist in the following sense. Regardless of which action player A uses in the stage game, by catching without restraint player B causes a loss of $\frac{1}{2}\mu_t$ to player A compared to the situation in which B were to catch with restraint. So, not only does B catch more by showing no restraint, the other agent suffers an immediate loss.

Now, we will specify μ_t. At stage $t \in \mathbb{N}$, the play is in state $s_t = \left(\rho_t^A, \rho_t^B\right)$, then the **normalized fish stock** is given by

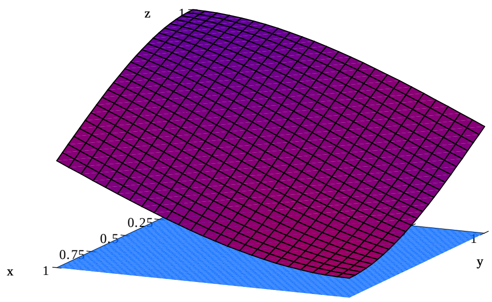
$$\mu_t \equiv 1 + 2\left(1 - \underline{m}\right)\left(\frac{\rho_t^A + \rho_t^B}{2}\right)^3 - 3\left(1 - \underline{m}\right)\left(\frac{\rho_t^A + \rho_t^B}{2}\right)^2, \qquad (2)$$

where $\underline{m} \in [0,1]$ represents minimal stock. We visualize (2) below taking $\underline{m} = \frac{1}{10}$.

So, the more frequent the players used the fine-mazed net in the stages leading up to stage t, the lower the number μ_t is. We have chosen an 'inverted' S-shaped curve, because this creates a rather convincing scenario.[3] For low values of $\rho_t^A + \rho_t^B$, the negative effects on the catches are quite small,

[3]S-shaped curves are widely used to model diffusion processes, learning processes, the spreading of infectuous diseases, and so on. We took parameters in a third-degree polynomial equation such that $v(0) = \overline{m} = 1$, $v(1) = \underline{m}$, and $v'(0) = v'(1) = 0$. The gist of the matter is that $v'(x) < 0$ for all $x \in (0,1)$.

as the fish stock recovers to near maximal levels; for intermediate values of $\rho_t^A + \rho_t^B$, the negative effects on the catches are more significant; for high values, the effects taper off again, as the small fish stock allows only rather small catches of ten percent of the original catches.



The function $\mu(x,y) = 1 + \frac{9}{5}\left(\frac{x+y}{2}\right)^3 - \frac{27}{10}\left(\frac{x+y}{2}\right)^2$ on the interval $[0,1]^2$.

If $\underline{m} = 1$, then we have the same stage game over and over again, since $\mu_t = 1$ for all stages t. This implies that we are in the standard repeated game framework. Decreasing the level of $\underline{m}$ slightly yields stage games that are quite similar. However, we are particularly interested in fishery games which induce social dilemmas and social traps. This places an upper bound on $\underline{m}$. In the next section, we intend to be more specific about this issue.

We have illustrated the effects of changes in μ_t on feasible stage-payoffs in Figure 1, the lower left hand quadrangle represents possible stage payoffs which are a mere ten percent of the possible stage payoffs represented by the upper right hand quadrangle.

3 Strategies and rewards

At stage t, both players know the current state and the history of play[4], i.e., the state visited and actions chosen at stage $u < t$ denoted by $\left(s_u, j_u^A, j_u^B\right)$. A **strategy** prescribes at all stages, for any state and history, a mixed action to be used by a player. The sets of all strategies for A respectively B will be denoted by $\mathcal{X}^A$ respectively $\mathcal{X}^B$, and $\mathcal{X} \equiv \mathcal{X}^A \times \mathcal{X}^B$. The payoff to player k, $k = A, B$, at stage t, is stochastic and depends on the strategy-pair $(\pi, \sigma) \in \mathcal{X}$; the **expected stage payoff** is denoted by $R_t^k(\pi, \sigma)$.

[4] An inability to observe actions chosen or to distinguish the current state, leads to a repeated game with incomplete information (see e.g., Hart [1985]).

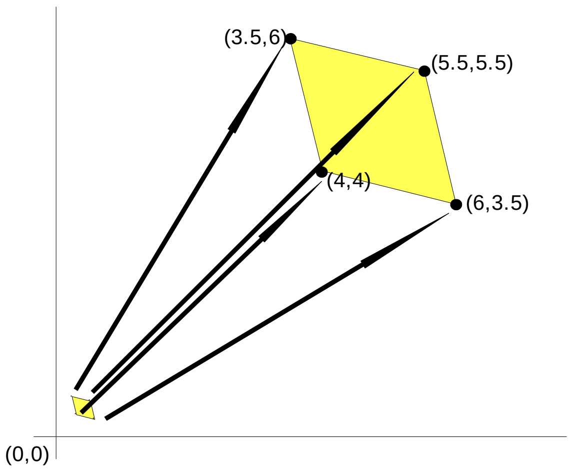


Figure 1: The feasible stage-payoffs of the Fishery Game for $\underline{m} = \frac{1}{10}$. The arrows indicate that the stage payoffs move upward if μ_t increases.

The players receive an infinite stream of stage payoffs during the play, and they are assumed to wish to maximize their average rewards. For a given pair of strategies (π, σ), player k's **average reward,** $k = A, B$, is given by $\gamma^k(\pi, \sigma) = \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} R_t^k(\pi, \sigma)$; $\gamma(\pi, \sigma) \equiv \left(\gamma^A(\pi, \sigma), \gamma^B(\pi, \sigma)\right)$.

It may be quite hard to determine the **set of feasible (average) rewards** F, directly. It is not uncommon in the analysis of repeated or stochastic games to limit the scope of strategies on the one hand, and to focus on rewards on the other. Here, we will do both, we focus on rewards from strategies which are pure and jointly convergent. Then, we extend our analysis to obtain more feasible rewards.

A strategy is **pure**, if at *each* stage a **pure action** is chosen, i.e., the action is chosen with probability 1. The set of pure strategies for player k is $\mathcal{P}^k$, and $\mathcal{P} \equiv \mathcal{P}^A \times \mathcal{P}^B$. The strategy pair $(\pi, \sigma) \in \mathcal{X}$ is **jointly convergent** if and only if $z \in \Delta^{m\times n}$ exists such that for all $\varepsilon > 0$:

$$\limsup_{t\to\infty} \Pr_{\pi,\sigma}\left[\left|\frac{\#\{j_u^A = i \text{ and } j_u^B = j \mid 1 \le u \le t\}}{t} - z_{ij}\right| \ge \varepsilon\right] = 0 \text{ for all } (i, j) \in J,$$

where $\Delta^{m\times n}$ denotes the set of all nonnegative $m \times n$-matrices such that the entries add up to 1, hence $z_{ij} \in [0, 1]$; $\Pr_{\pi,\sigma}$ denotes the probability under strategy-pair (π, σ). $\mathcal{JC}$ denotes the set of jointly-convergent strategy pairs. Under a pair of jointly-convergent strategies, the relative frequency of action pair $(i, j) \in J$ converges with probability 1 to z_{ij} in the terminology of Billingsley [1986, p.274]. Moreover, the empirical distribution of the past play by A under such a pair of strategies converges with probability 1 to the vector given by the row-sums of the matrix z. Hence, ρ_t^A converges with probability 1 to Z^A, i.e., the sum of the first row of the matrix z. Similar remarks hold with respect to the other player.

The **set of jointly-convergent pure-strategy rewards** is given by

$$P^{\mathcal{JC}} \equiv cl \left\{ \left(x^1, x^2\right) \in \mathbb{R}^2 | \ \exists_{(\pi,\sigma)\in\mathcal{P}\cap\mathcal{JC}} : \left(\gamma^k\left(\pi,\sigma\right), \gamma^k\left(\pi,\sigma\right)\right) = \left(x^1, x^2\right) \right\},$$

where $cl\ S$ is the closure of the set S. The interpretation of this definition is that for any pair of rewards in this set, we can find a pair of jointly-convergent pure strategies that yield rewards arbitrarily close to the original pair of rewards. $P^{\mathcal{JC}}$ can be determined rather conveniently, as we will show now. With respect to jointly-convergent strategies, Eq. (2) and the arguments presented imply that

$$\lim_{t\to\infty} \mu_t\left(a_{ij}, b_{ij}\right) = \left(1 + 2\left(1 - \underline{m}\right) Z^3 - 3\left(1 - \underline{m}\right) Z^2\right)\left(a_{ij}, b_{ij}\right),$$

where $Z \equiv \frac{Z^A + Z^B}{2}$. So, the bi-matrices $\left(A\left(\mu_t\right), B\left(\mu_t\right)\right)$ 'converge' in the long run, too.

Let $\varphi(z) \equiv \left(1 + 2\left(1 - \underline{m}\right) Z^3 - 3\left(1 - \underline{m}\right) Z^2\right) \sum_{(i,j)\in J} z_{ij}\left(a_{ij}, b_{ij}\right)$. The interpretation of $\varphi(z)$ is that under jointly-convergent strategy pair (π, σ) the relative frequency of action pair $(i, j) \in J$ being chosen is z_{ij} and each time this occurs the players receive $\left(a_{ij}\mu_t, b_{ij}\mu_t\right)$ in the long run. Hence, the players receive an average amount of $\varphi(z)$. So, $\gamma\left(\pi, \sigma\right) = \varphi(z)$.

The following result was proven in Joosten *et al.* [2003] for FD-games. Less general ideas had been around earlier for the analysis of repeated games with vanishing actions (cf., Joosten [1996, 2001], Schoenmakers *et al.* [2002]). Let $CP^{\mathcal{JC}}$ denote the convex hull of $P^{\mathcal{JC}}$.

Theorem 1 *(Joosten, Brenner & Witt [2003]) For any FD-game, we have* $P^{\mathcal{JC}} = \bigcup_{z\in\Delta^{m\times n}} \varphi(z)$. *Moreover, each pair of rewards in* $CP^{\mathcal{JC}}$ *is feasible.*

The set of **Pareto-efficient outcomes in** $CP^{\mathcal{JC}}$ is given by

$$PE \equiv \left\{ \left(x^A, x^B\right) \in CP^{\mathcal{JC}} | \nexists \left(y^A, y^B\right) \in CP^{\mathcal{JC}} : \left(y^A, y^B\right) > \left(x^A, x^B\right) \right\}.$$

The inequality sign here refers to both components. We refer to Figure 2 for an illustration. To give the intuition on how to construct strategies which yield a reward in $CP^{\mathcal{JC}} \backslash P^{\mathcal{JC}}$, we provide the example below.

Example 2 *To illustrate how to obtain rewards in* $CP^{\mathcal{JC}} \backslash P^{\mathcal{JC}}$, *consider the following where we take* $\underline{m} = \frac{1}{10}$.
Let for $i \in \{1, 2\}$, $\pi^i = \sigma^i = (i, i, i, ...)$. *Then,* $\gamma\left(\pi^1, \sigma^1\right) = (0.55; 0.55) \in P^{\mathcal{JC}}$ *and* $\gamma\left(\pi^1, \sigma^2\right) = (1.925; 3.3) \in P^{\mathcal{JC}}$ *because both pairs of strategies are clearly jointly-convergent. However, it can be confirmed in Figure 2 that no other convex combination of both rewards belongs to* $P^{\mathcal{JC}}$.

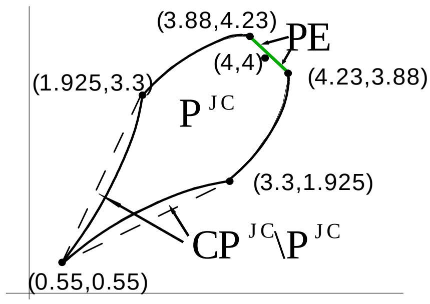


Figure 2: The shaded area denotes $P^{\mathcal{JC}}$; $CP^{\mathcal{JC}}$, the convex hull of $P^{\mathcal{JC}}$, can be obtained using special types of mixed strategies; clearly, $(4,4) \notin PE$.

We will now show that a pair of strategies (π,σ) *exists for which* $\gamma(\pi,\sigma) = \frac{3}{4}\gamma\left(\pi^1,\sigma^1\right)+\frac{1}{4}\gamma\left(\pi^1,\sigma^2\right)$. *Define* (π,σ) *by:*

$$\begin{array}{ll} \pi_t=\sigma_t=\left(\frac{1}{2},\frac{1}{2}\right) & \textit{for } t=1,2, \textbf{\textit{ followed by:}} \\ (\pi_t,\sigma_t)=\left(\pi_t^1,\sigma_t^2\right) & \textit{if } \left(\left(j_1^A,j_1^B\right),\left(j_2^A,j_2^B\right)\right)=\left((i,i),(j,j)\right),\ i,j\in\{1,2\} \\ (\pi_t,\sigma_t)=\left(\pi_t^1,\sigma_t^1\right) & \textit{otherwise.} \end{array}$$

So, at the first two stages both players randomize with equal probability on both actions. Then, sixteen outcomes $\left(\left(j_1^A,j_1^B\right),\left(j_2^A,j_2^B\right)\right)$ *are possible, each one being equally likely; four outcomes induce play according to* $\left(\pi^1,\sigma^2\right)$ *for the remainder, twelve induce play according to* $\left(\pi^1,\sigma^1\right)$. *The long run average stage payoffs converge to* $(1.925;3.3)$ *in the first case, and to* $(0.55;0.55)$ *in the second one. This means that* $\gamma(\pi,\sigma)=\frac{4}{16}(1.925;3.3)+\frac{12}{16}(0.55;0.55)=\frac{1}{4}\gamma\left(\pi^1,\sigma^2\right)+\frac{3}{4}\gamma\left(\pi^1,\sigma^1\right)$. *Randomizing during the first two stages and choosing appropriate 'reactions' for the remainder of the play similar to the one described, yields all convex combinations of* $\gamma\left(\pi^1,\sigma^1\right)$ *and* $\gamma\left(\pi^1,\sigma^2\right)$ *which are multiples of* $\frac{1}{4}$. *This is generalizable to larger randomization periods* T *yielding all multiples of* $\frac{1}{2}^T$. ■

Figure 3 depicts the sets of jointly-convergent pure-strategy rewards for different values of $\underline{m}$. For the case depicted on the left, there exists no social trap problem at all. For a social trap to occur, we must have that the 'no-restraint' reward, i.e., $\frac{11}{2}\underline{m}$, is lower than the 'perfect restraint' reward, i.e., 4. This imposes the restriction $\underline{m} \leq \frac{8}{11}$. The other case shows that the 'no-restraint' rewards are $(3,3)$, hence there exists only a fairly moderate social trap problem. In the remainder, we will use the following.

Remark 1 *We assume that the 'no-restraint' outcome gives at most half the rewards associated with the 'perfect restraint' outcome.*

With respect to the minimal level this implies $\underline{m} \leq \frac{4}{11}$. This also serves to avoid several technical difficulties for the remaining analysis which have some appeal, but do not necessarily contribute to added conceptual insights.

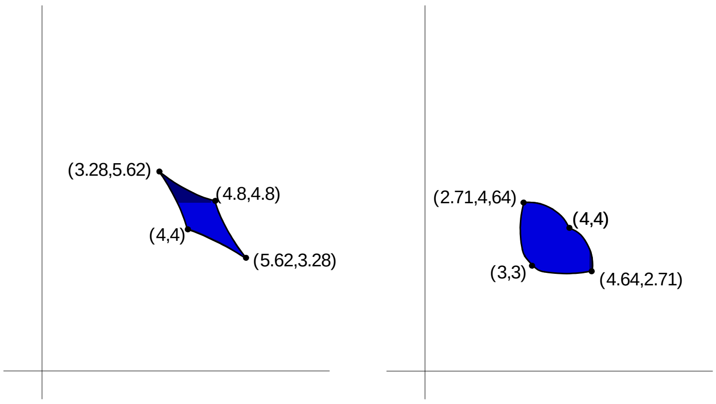


Figure 3: On the left-hand side, we have the jointly-convergent pure-strategy rewards for $\underline{m} = \frac{96}{110}$ and on the right-hand side for $\underline{m} = \frac{6}{11}$.

4 Threats and equilibria

The strategy pair (π^*, σ^*) is an **equilibrium**, if no player can improve by unilateral deviation, i.e., $\gamma^A(\pi^*, \sigma^*) \geq \gamma^A(\pi, \sigma^*)$, $\gamma^B(\pi^*, \sigma^*) \geq \gamma^B(\pi^*, \sigma)$ for all $\pi \in \mathcal{X}^A, \sigma \in \mathcal{X}^B$. An equilibrium is called **subgame perfect** if for each possible state and possible history (even unreached states and histories) the subsequent play corresponds to an equilibrium, i.e., no player can improve by deviating unilaterally from then on. In the construction of equilibria for repeated games, 'threats' play an important role. A threat specifies the conditions under which one player will punish the other, as well as the subsequent measures.

We call $v = (v^A, v^B)$ the **threat point**, where $v^A = \min_{\sigma \in \mathcal{X}^B} \max_{\pi \in \mathcal{X}^A} \gamma^A(\pi, \sigma)$, and $v^B = \min_{\pi \in \mathcal{X}^A} \max_{\sigma \in \mathcal{X}^B} \gamma^B(\pi, \sigma)$. So, v^A is the highest amount A can get if B tries to minimize his average payoffs. Under a pair of **individually rational** (feasible) rewards each player receives at least the threat-point reward. We have the following result for the Fish War.

Lemma 3 *Under Remark 1, we have* $v = (1 + \underline{m}) \left(\frac{7}{4}, \frac{7}{4}\right)$.

Example 4 *Consider the pair of strategies in which A fishes alternately with fine-mazed nets and with large-mazed nets, and B alternating vice versa. Then,* $\pi = (1,2,1,2,1,2,1,2,...)$ *and* $\sigma = (2,1,2,1,2,1,2,1,...)$. *This implies that* $\rho_t^A = \rho_t^B = \frac{1}{2}$ *in the long run. So, for* $\underline{m} = \frac{1}{10}$ *we have* $\lim_{t\to\infty} \mu_t = \frac{11}{20}$. *Thus, B's long run stage payoffs are alternately* $\frac{7}{2}\mu_t = \frac{77}{40}$ *and* $6\mu_t = \frac{66}{20}$, *the average stage payoffs converge to* $\frac{1}{2}\left(\frac{77}{40} + \frac{66}{20}\right) = \frac{209}{80} = 2.612\,5$. *A similar statement holds for A. So,* $\gamma(\pi, \sigma) = (2.612\,5, 2.612\,5)$
Suppose A were to deviate unilaterally, and fish with fine-mazed nets only once every four stages exactly when B does not fish with fine-mazed nets. So, $\widetilde{\pi} = (2,2,1,2,2,2,1,2,...)$ *and* $\sigma = (2,1,2,1,2,1,2,1,...)$. *Then,* $\rho_t^A = \frac{1}{4}$, $\rho_t^B = \frac{1}{2}$ *and* $\mu_t = 0.715\,23$ *in the long run. Then A receives* $\frac{7}{2}\mu_t$ *twice, and* $6\mu_t$ *and* $4\mu_t$ *once each in every four stages. This leads to average stage payoffs of* $3.039\,7$. *As A can improve unilaterally against B's strategy σ, (π, σ) is not a Nash equilibrium.*
The threat point is $v = \left(\frac{77}{40}, \frac{77}{40}\right) = (1.925, 1.925)$. *Next, consider the following pair of strategies in which A plays according to sequence* $(1,2,1,2,1,2,1,2,...)$ *and B according to sequence* $(2,1,2,1,2,1,2,1,...)$ *as long as both players stick to their respective sequences. If one player deviates from this course of action, i.e., the sequence given above, the other player is never to fish with restraint again, i.e., then the punisher uses the action 1 from then on.*
Clearly, this pair of strategies leads to exactly the same sequence of play as the strategy pair $\left(\pi^A, \sigma^B\right)$ *does, inducing rewards* $\frac{209}{80}$ *to both players. However, if A were to deviate only once against the strategy employed by B in this case, then B would 'punish' this deviation by never showing restraint again. So, A's reward is at most* $\frac{77}{40}$. *Hence, A can not improve unilaterally. A similar argument holds if B were to deviate unilaterally. The conditions for a Nash equilibrium are therefore fulfilled by this new pair of strategies, where for (π, σ) they were not.* ■

To present the general idea of the result of Joosten *et al.* [2003] to come, we adopt terms from Hart [1985] and Forges [1986]. First, there is a 'master plan' which is followed by each player as long as the other does too; then there are 'punishments' which come into effect if a deviation from the master plan occurs. The master plan is a sequence of 'intra-play communications' between the players, the purpose of which is to decide by which equilibrium the play is to continue. The outcome of the communication period is determined by a 'jointly controlled lottery', i.e., at each stage of the communication period the players randomize with equal probability on the first two actions; at the end of the communication period one sequence of pairs of action choices materializes. We have illustrated in Example 2 how such a jointly-controlled lottery can be accomplished and in effect this method is the basic tool in the formal proofs given in Joosten *et al.* [2003].

Detection of deviation from the master plan *after* the communication period is easy as both players use pure actions on the equilibrium path from then on. Deviation in the communication period by using an *alternative randomization* on the first two actions is impossible to detect. However, it can be shown that no alternative unilateral randomization yields a higher reward. So, the outcome of the procedure is an equilibrium. For more details, we refer to Joosten *et al.* [2003]. We restate here the major result which applies to general games with frequency-dependent stage payoffs and therefore clearly has implications for our FD-Fish War.

Theorem 5 *(Joosten, Brenner & Witt [2003]) Each pair of rewards in the convex hull of all individually-rational pure-strategy rewards can be supported by an equilibrium. Moreover, each pair of rewards in the convex hull of all pure-strategy rewards giving each player strictly more than the threat-point reward, can be supported by a subgame-perfect equilibrium.*

The following corollary is illustrated in Figure 4.

Corollary 6 *Let* $E' = \{(x, y) \in P^{\mathcal{JC}} | \ (x, y) \geq v\}$ *be the set of all individually rational jointly-convergent pure strategy rewards in the FD-Fish War. Then, each pair of rewards in the set* E' *can be supported by an equilibrium. Moreover, all rewards in* E' *giving* A *strictly more than* v^A *and* B *strictly more than* v^B *can be supported by a subgame-perfect equilibrium.*

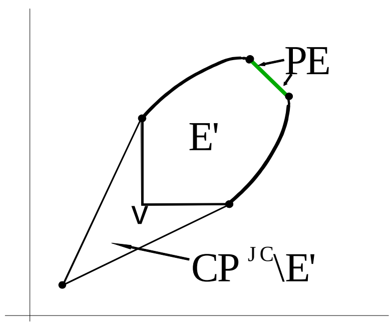


Figure 4: The shaded area represents equilibrium rewards.

5 Discussion and ideas for further research

The final theorem of the previous section is of course a Folk Theorem. Two aspects are debatable and in fact draw the bulk of the critical remarks. First, there exists a multitude of equilibria in general and it is not obvious which equilibria should be chosen. Second, punishments which are crucial to many of the equilibria found, can be costly and quite unforgiving (see e.g., Gintis [2001]). By these remarks we mean that the punisher often has to disregard his own interests in the interest of punishing his opponent for undesired behavior; furthermore, often grim trigger strategies are used in which even a single deviation will trigger an eternal punishment.

On the positive side, Osborne & Rubinstein [1994] point out that equilibria of the infinitely repeated game exist which are Pareto-superior to any equilibrium of the associated one-shot game. The Pareto-efficient equilibrium rewards within the set of jointly-convergent pure-strategy rewards are denoted by PE in Figure 4. What is striking is the Pareto-efficient rewards can only be obtained if the Pareto-inferior action in the stage games is played with a long-run relative frequency significantly higher than zero by both players. In the example used for expository purposes throughout this paper, we find that Pareto-efficient equilibria

- yield combined rewards which are slightly (1.4%) higher than the combined 'perfect restraint' equilibrium rewards, yet *more than seven times* the combined 'no-restraint' rewards;
- induce play in which both players simultaneously show restraint for approximately 85.6% of the stages, in the remaining stages, only one player shows restraint.

The consistency requirement of Güth *et al.* [1991] requires that in identical subgames the players use exactly the same 'substrategies'. If all subgames are equal to the original game, then the players must choose the same action in all stages, i.e., they must employ a stationary strategy. This excludes the option of using history-dependent strategies, and helps to reduce the number of equilibria. However, there is a tension between consistency and Pareto-efficiency in general. For the Prisoners' Dilemma consistency implies that a unique equilibrium remains namely the one in which the players always defect; all other equilibrium rewards found with the Folk Theorem are Pareto-superior.

The richness of the strategy-space allows us to accommodate however, several of these objections connected to equilibria involving threats. We now show that we can construct subgame-perfect equilibria which in case of unilateral deviations are 'forgiving', i.e., they allow (not too many) deviations, and if a deviator is punished the 'punisher' is better off afterwards.

Theorem 7 *For any pair $(a, b) \in PE \cup int\ E'$ a subgame-perfect equilibrium (π, σ) exists yielding rewards (a, b) which induces play such that*

- *if A deviates from the equilibrium path 'too often', then play proceeds according to an equilibrium such that A receives strictly less than the amount a but more than v^A and B receives at least b;*

- *if B deviates from the equilibrium path 'too often', then play proceeds according to an equilibrium such that B receives strictly less than b but more than v^B and A receives at least a.*

The 'too often'-s are well specified in the proof. Again, this approach does not significantly reduce the set of equilibria.

One of our concessions to reality involved what the players know. In real-world common-pool systems, actions of other agents are often not directly observable. Several field studies show a significant positive effect of monitoring on the propensity to adhere to a socially acceptable course of action, e.g., to show restraint in our model. The same is true about the ability to punish certain socially undesirable behavior (see e.g., Ostrom [1994]). Note that though theoretically and practically possible, hardly any field study of real-life common-pool resource dilemmas reports the use of anything comparable to grim-trigger strategies (Ostrom *et al.* [1994], Ostrom [2004]).

An alternative evaluation criterion used in economics involves discounting. For strategy-pair (π, σ) and discount factor $\beta \in [0, 1)$, the **discounted reward** of player k, is given by $\gamma_\beta^k(\pi, \sigma) = (1 - \beta) \sum_{t=1}^{\infty} \beta^{t-1} R_t^k(\pi, \sigma)$. The parameter β represents the time-preferences of the agents; the part before summation sign is a normalization term. A small β implies that the agents are impatient, i.e., they attribute a lot of weight to the near future; a high value of the same parameter indicates patience on the part of the agents.

Our preference for the average reward criterion comes from the following considerations. First, there is a fairly consistent mapping of what the common-pool resource will yield in the long run and how the agents evaluate their revenues from it. With regard to discounting, the mappings (plural!) are far more intricate. Second, discounting in environmental issues and in social dilemmas is rather problematic. Under discounting, any immediate advantage is preferred to any environmental, economic or social catastrophe *sufficiently far away in the future*. The actual discounting factor is irrelevant in this critique. So, a social trap is actually far more likely under discounting than under the criterion which we used.

Brennan *et al.* [2004] advocate an approach in which modelers (loosely speaking) confine themselves to those models in which first there exists a structural correspondence between 'model and reality', and second a 'continuity' can be demonstrated between 'ideal' and the 'real'. Here, this would involve showing that equilibria for the limiting average reward criterion are also equilibria for the limiting case, i.e., $\beta \uparrow 1$. Here, however, the limiting case does not seem the problem, rather that a very large lower bound on β exists for which the desired properties hold.

Our modeling was guided by a desire to reproduce salient features of strategic interaction in fisheries using simple mathematical relations. Experienced researchers in this field may feel that we build too naive models and have ignored relevant literature. However, our approach does not rely on the simple mathematical relationships chosen here, a wide range of alternatives can be analyzed in exactly the same manner. Also, we intend to incorporate more advanced ideas from the renewable resources literature in future projects. The natural extensions to more agents in the game and addition of more complex strategic interaction between those agents, must also be reserved for future research.

6 Appendix

Proof of Lemma 3. We only prove the case for v^A. Our claim is that (π, σ) given by $\pi = (2, 2, 2, ...)$ and $\sigma = (1, 1, 1, ...)$ is the pair of strategies such that σ minimizes A's maximal reward, and π is A's best reply. It can be confirmed that $\frac{\partial \mu_t}{\rho_t^A}, \frac{\partial \mu_t}{\rho_t^B} \leq 0$ with strict inequality if $\rho_t^A + \rho_t^B \neq 2$. Hence, player B can minimize μ_t by using σ taking the strategy of his opponent as given. Moreover, if B uses action 1 in a stage game, then the payoffs in the stage game are minimized. So, the long run stage payoffs of player A are minimized by player B by using σ.
Similarly, player A maximizes μ_t by using π taking the strategy of his opponent as given. However, the conclusion regarding his maximization problem is not as easily answered as B's minimization problem. Denote $R_t^A(x, 1) =$

$$\left(\frac{11}{2}x + \frac{7}{2}(1-x)\right)\left(1 + 2(1-\underline{m})\left(\frac{x+1}{2}\right)^3 - 3(1-\underline{m})\left(\frac{x+1}{2}\right)^2\right)$$

then it can be confirmed that $R_t^A(x, 1) \leq R_t^A(0, 1)$ for all $x \in [0, 1]$, $\underline{m} \in \left[0, \frac{4}{11}\right]$. This means that whichever strategy player A uses, the long run stage payoff will be at most $R_t^A(0, 1)$. Hence, $\gamma(\widetilde{\pi}, \sigma) \leq R_t^A(0, 1) = \gamma(\pi, \sigma) = v^A$ for all $\widetilde{\pi}$. ■

Proof of Theorem 7. Let (x, y) be a pair of rewards in the interior of E'. Then, numbers $\underline{x} < x \leq \overline{x}$ and $\underline{y} < y \leq \overline{y}$ exist such $(\underline{x}, \overline{y}), (\overline{x}, \underline{y}) \in E'$. So, equilibrium strategies (π^x, σ^y), $\left(\pi^{\underline{x}}, \sigma^{\overline{y}}\right)$, $\left(\pi^{\overline{x}}, \sigma^{\underline{y}}\right)$ exist such that $\gamma(\pi^x, \sigma^y) = (x, y)$, $\gamma\left(\pi^{\underline{x}}, \sigma^{\overline{y}}\right) = (\underline{x}, \overline{y})$, $\gamma\left(\pi^{\overline{x}}, \sigma^{\underline{y}}\right) = \left(\overline{x}, \underline{y}\right)$. Let $T^* \geq 0$ denote the length of the communication period of strategy pair (π^x, σ^y). For $T > 1$, let $AD_{\pi^x}^A(T) = \frac{\#\{j_t^A \neq \pi_t^x | t \leq T\}}{T}$, $AD_{\sigma^y}^B(T) = \frac{\#\{j_t^B \neq \sigma_t^y | t \leq T\}}{T}$. Define (π, σ) as follows:

$$\begin{array}{lll}
\pi_t = \pi_{t-T'}^{\underline{x}} & \sigma_t = \sigma_{t-T'}^{\overline{y}} & \left\{\begin{array}{l}\text{if for some } T' > T^*\text{: } AD_{\pi^x}^A(T') > 1/\sqrt{T'} \\ \text{and } AD_{\pi^x}^A(T') \geq AD_{\sigma^y}^B(T')\end{array}\right. \\
\pi_t = \pi_t^{\overline{x}} & \sigma_t = \sigma_t^{\underline{y}} & \left\{\begin{array}{l}\text{if for some } T'' > T^*\text{: } AD_{\sigma^y}^B(T'') > 1/\sqrt{T''} \\ \text{and } AD_{\sigma^y}^B(T'') > AD_{\pi^x}^A(T'')\end{array}\right. \\
\pi_t = \pi_t^x & \sigma_t = \sigma_t^y & \text{otherwise.}
\end{array}$$

Then, $\gamma(\pi, \sigma) = \gamma(\pi^x, \sigma^y) = (x, y)$, because in the long run deviations from the equilibrium path of (π^x, σ^y) go to zero in relative frequency. Now, (π, σ) is an equilibrium because if, e.g., B deviates more than A does such that for some T'' :

$AD^{B}_{\sigma^y}(T'') > 1/\sqrt{T''}$, then the play continues according to equilibrium $\left(\pi^{\overline{x}}, \sigma^{\underline{y}}\right)$. B gets $\underline{y} < y$ in that case, while A receives $\overline{x} \geq x$. A similar statement holds for A as well. Hence, neither player can improve his rewards by deviating unilaterally. ■

7 References

Amir R, 2003, Stochastic games in economics and related fields: an overview, in: "Stochastic Games and Applications" (A Neyman & S Sorin, eds.), *NATO Advanced Study Institute, Series D*, Kluwer, Dordrecht, pp. 455-470.

Billingsley P, 1986, "Probability and Measure", John Wiley & Sons, New York, NY.

Brenner T, & U Witt, 2003, Melioration learning in games with constant and frequency-dependent payoffs, *J Econ Behav Organ* **50**, 429-448.

Brennan G, W Güth & H Kliemt, 1994, Approximate truth in economic modelling, Discussion Papers on Strategic Interaction #38-2004, Max Planck Institute, Jena, Germany.

Dawes RM, 1975, Formal models of dilemmas in social decision-making, in: "Human Judgment and Decision Processes" (MF Kaplan & S Schwartz, eds.), Academic Press, New York, NY, pp. 87-108.

Dawes RM, 1980, Social dilemmas, *Ann Rev Psychology* **21**, 169-193.

Dockner E, S Jørgensen, N Van Long, & G Sorger, "Differential Games in Economics and Management Science", Cambridge University Press, Cambridge, UK.

Forges F, 1986, An approach to communication equilibria, *Econometrica* **54**, 1375-1385.

Gintis H, 2001, "Game Theory Evolving", Princeton University Press, Princeton, NJ.

Güth W, W Leininger & G Stephan, 1991, On supergames and Folk Theorems: A conceptual discussion, in "Game Equilibrium Models. Morals, Methods,and Markets" (R Selten, ed.), Springer, Berlin, pp. 56-70.

Hardin G, 1968, The tragedy of the commons, *Science* **162**, 1243-1248.

Hanley N, JF Shogren & B White, 1997, "Environmental Economics in Theory and Practice", MacMillan Press Ltd., Basingstoke.

Hart S, 1985, Nonzero-sum two-person repeated games with incomplete information, *Math Oper Res* **10**, 117-153.

Herrnstein RJ, 1997, "The Matching Law: Papers in Psychology and Economics", Harvard University Press, Cambridge, MA.

Joosten R, 1996, "Dynamics, Equilibria, and Values", Ph.D.-thesis, Maastricht University.

Joosten R, 2001, A note on repeated games with vanishing actions, forthcoming in *Int Game Theory Rev.*

Joosten R, T Brenner & U Witt, 2003, Games with frequency-dependent stage payoffs, *Int J Game Theory* **31**, 609-620.

Komorita SS & CD Parks, 1994, "Social Dilemmas", Brown & Benchmark's Social Psychology Series, Brown & Benchmark, Madison WI.
Levhari & Mirman, 1980, The great fish war: An example using a dynamic Cournot-Nash solution, *Bell J Econ* **11**, 322-334.
Osborne MJ, & A Rubinstein, 1994, "A Course in Game Theory", MIT Press, Cambridge, MA.
Ostrom E, 2000, Private and Common Property Rights, in: "Encyclopedia of Law and Economics", *Vol.* **II**, B.Bouckaert & G. De Geest (eds.), Edward Elgar, Cheltenham, UK, pp. 322-379.
Ostrom E, 2004, The working parts of rules and how they evolve over time, *Papers on Economics & Evolution* **#0404**, Max Planck Institute Jena.
Ostrom E, R Gardner, & J Walker, 1994, "Rules, Games, and Common-Pool Resources", The University of Michigan Press, Ann Arbor, MI.
Platt J, 1973, Social traps, *American Psychologist* **28**, 641-651.
Schoenmakers G, J Flesch & F Thuijsman, 2002, Coordination games with vanishing actions, *Int Game Theory Rev* **4**, 119-126.
Shapley L, 1953, Stochastic games, *Proceedings of the National Academy of Sciences U.S.A.* **39**, 1095-1100.
Shy O, 1995, "Industrial Organization: Theory and Applications", The MIT Press, Cambridge, MA.
Thuijsman F, 1992, "Optimality and Equilibria in Stochastic Games", *CWI-Tract* **82**, Centre for Mathematics and Computer Science, Amsterdam.
Tolwinski B, A Haurie, G Leitmann, 1986, Cooperative equilibria in differential games, *J Math Anal Appl* **119**, 182-202.
Vieille N, 2000a, Two-player stochastic games I: A reduction, *Israel J Math* **119**, 55-91.
Viellie N, 2000b, Two-player stochastic games II: The case of recursive games, *Israel J Math* **119**, 93-126.
Vrieze OJ, 1987, "Stochastic games with Finite State and Action Spaces", *CWI-Tract* **33**, Centre for Mathematics and Computer Science, Amsterdam.