1024

Stochastic games with endogenous transitions

by

Reinoud Joosten
Robin Meijboom

Stochastic games with endogenous transitions

Reinoud Joosten,[*] Robin Meijboom[†]

November 25, 2010

Abstract

We introduce a stochastic game in which transition probabilities depend on the history of the play, i.e., the players' past action choices. To solve this new type of game under the limiting average reward criterion, we determine the set of jointly-convergent pure-strategy rewards which can be supported by equilibria involving threats.

We examine the following setting for motivational and expository purposes. Each period, two agents exploiting a fishery choose between catching with restraint or without. The fish stock is in either of two states, *High* or *Low*, and in the latter each action pair yields lower payoffs. Restraint is harmless to the fish, but it is a dominated strategy in each stage game. Absence of restraint damages the resource, i.e., the less restraint the agents show, the higher the probablities that *Low* occurs at the next stage of the play. This state may even become 'absorbing', i.e., transitions to *High* become impossible.

JEL-Codes C72, C73, Q22, Q57

Keywords Stochastic games, endogenous transitions, limiting average rewards, equilibria, common pool resource dilemma

1 Introduction

The main motivation of this paper is to incorporate certain empirical phenomena of a stochastic nature into Small Fish Wars, i.e., a class of models on the exploitation of a renewable natural resource featuring strategic interaction by multiple agents (cf., Joosten [2007a,b,c, 2010a]). One extension of these Small Fish Wars leads to a new type of stochastic game.[1]

---

[*]Corresponding author: FELab, School of Management & Governance, University of Twente, POB 217, 7500 AE Enschede, The Netherlands: r.a.m.g.joosten@utwente.nl

[†]Deloitte Accountants BV, Enterprise Risk Services, Laan van Kronenburg 2, 1183 AS Amstelveen, The Netherlands.

[1]Stochastic games were introduced by Shapley [1953]; for overviews and results see e.g., Vrieze [1987], Thuijsman [1992], Flesch [1998] and Vieille [2000a,b]. Amir [2003] connects stochastic and difference games in economic applications.

A *Small Fish War*[2] (Joosten [2007a]) has the following set-up. Two agents possess the fishing rights to a body of water, and they have essentially two options, to fish with or without restraint. Restraint may take various forms, e.g., catching seasons, quantities caught, catching methods, or technologies, e.g., boats, nets, allowed in catching. Essential is that unrestrained fishing yields a higher immediate catch, but if continued, it may damage the resource and therefore lead to decreasing catches in the future. Restrained fishing by both agents is sustainable in the long run.

In the standard version, agents wish to maximize their average catches over an infinite time-horizon. In such a setting, a 'tragedy of the commons' does not seem inevitable, as Pareto-efficient outcomes can be sustained by subgame perfect equilibria. Moreover, all equilibria yield rewards which are well-above those connected to ruthless exploitation of the resource, because a sizeable fish stock is preserved as time continues. The intuition behind this is that catches, hence the stage payoffs of the agents, are linked to the level of the fish stocks and are therefore rather high on the long run. This results in rewards being rather high as well. Ruthless exploitation of the resource is good for immediate profits, yet profits will decline as time continues. As a result their averages may become rather low.

In this paper we add randomness to a Small Fish War in three variants. First, we introduce a simple form of stochasticity into the most simplified version of the Small Fish War model. We take two states, *High* and *Low*, and assume that there exist exogenously given transition probabilities to move from one to the other, which do not depend on the actions taken by the agents. This set-up determines also the long run transitions from *High* to *Low* and vice versa and hence, the proportion of time the play spends in either state, as well. Examples of this kind of stochasticity might be the influence of weather, climate, water temperatures, i.e., events (re)occurring independently from actions the agents take.

In the second variant, the exogenously given transition probabilities depends on the current state and actions chosen by the agents. Here, the proportion of time the play spends in either state, may vary depending on the actual play. In that sense, although the transition probabilities are fixed, the agents influence the long run distribution of the play over the two states. This type of stochasticity arises if the propensy of the system moving to *High* depends on the current level of restraint on the part of the agents, as well as on the current state. For instance, the same lack of restraint may be more damaging to the resource if it is scarce than if it is abundant.

The third innovation is more fundamental. We introduce endogenous stochastic variation, i.e., transition probabilities are influenced by past catching behavior of the agents. This extension is motivated by certain hysteresis ef-

---

[2]A word play on the title of the paper by Levhari & Mirman [1980] showing that strategic interaction in a fishery may induce a 'tragedy of the commons' (Hardin [1968]).

fects called poaching pits[3] (e.g., Bulte [2003], Courchamp *et al.* [2006], Hall *et al.* [2008]). We proceed with basically the same set-up as in the previous variants, but the more heavily the agents exploit the resource as time continues, the worse its condition becomes. As a consequence, the probability of the system to move to *High* decreases in time in each state and for each action combination. So, the play occurs more often in *Low* and therefore, the catches of the agents will be low for a higher proportion of time. To solve this new kind of game, we draw inspiration from earlier work on Small Fish Wars, but also from Joosten, Peters & Thuijsman [1995], Joosten [1996,2005], Thuijsman & Vrieze [1998], Joosten, Brenner & Witt [2003], Schoenmakers *et al.* [2002] and Schoenmakers [2004].

For the setting to be presented and analyzed, the agents wish to maximize their long term average catches as in standard Small Fish Wars. In our analysis we use an approach rather similar to the one in Joosten *et al.* [2003] and provide the necessary proofs to validate these procedures in this new setting. First, we prove that the rewards for any pair of jointly-convergent pure strategies are easily established. Then, we determine the set of jointly-convergent pure-strategy rewards. A more complex issue is then to find for each player the threat point reward, i.e., the highest amount this player can obtain if his opponent tries to minimize his rewards. We finally obtain a large set of rewards which can be supported by equilibria using threats.

We were able to reproduce effects similar to the ones associated to a poaching pit by *Low* (temporarily) becoming an absorbing state dependent on the behavior of the agents. By a temporarily absorbing state we mean that all transition probabilities from this state become zero temporarily. If the agents keep overexploiting the resource, this situation will not change, i.e., the transition probabilities mentioned remain zero. If the agents were to adopt behavior permitting recovery of the resource, it may take some time before transition probabilities to *High* become nonzero again.

In the model examined for expository purposes, the proportion of time spent in such a poaching pit goes to zero in the long run in each subgame-perfect equilibrium behavior. Exactly one equilibrium, no restraint, can be associated to the poaching pit and, as mentioned, all other equilibria Pareto-dominate it. To put it differently, no subgame perfect equilibrium can be associated with the poaching pit and precisely one equilibrium can. It is also surprising that all jointly-convergent pure-strategies inducing the poaching pit, yield lower combined rewards than no-restraint.

Next, we introduce the two models with exogenous and semi-exogenous stochasticity. These models are fairly well-known in the analysis of stochastic games. In Section 3, we focus on strategies and rewards in a very general

---

[3]One characteristic of a poaching pit is that if the system is forced into critically low densities by overexploitation, it may take a rather long time before the resource recovers after a program of rehabilitation is adopted.

sense in order to prepare for the sequel. In Section 4, we present our model with completely endogenous transition probabilities. Section 5 continues the analysis for the models introduced in order to present sets of equilibrium rewards. Section 6 concludes.

2 (Semi-)exogenous variability in catch sizes

A *Small Fish War* is played by players A and B at discrete moments in time called stages. Each player has two actions and at each stage $t \in \mathbb{N}$ the players independently and simultaneously choose an action. Action 1 denotes the action for which some restriction exists allowing the resource to recover, e.g., catching with wide-mazed nets or catching a low quantity. Action 2 for player A (B) denotes the action without or with very little restraint, e.g., catching with fine-mazed net or catching a high quantity. In the following subsections, we treat two extensions of the Small Fish War model staying within the framework offered by stochastic game theory.

2.1 Exogenous variability

For the sake of simplicity we assume catches only to vary due to stochastic shocks, which we model by means of a stochastic game with two states, *High* and *Low*. Let the payoff matrices in the two states be given by

$$
H = \left[\begin{pmatrix} \theta_1 \\ p'_H \\ \theta_3 \\ p'_H \end{pmatrix} \begin{pmatrix} \theta_2 \\ p'_H \\ \theta_4 \\ p'_H \end{pmatrix} \right] \text{ and } L = \left[\begin{pmatrix} \theta_5 \\ p'_L \\ \theta_7 \\ p'_L \end{pmatrix} \begin{pmatrix} \theta_6 \\ p'_L \\ \theta_8 \\ p'_L \end{pmatrix} \right].
$$

Here, subscript H (L) indicates the payoff matrix to be associated with state *High* (*Low*). The catches in *Low* are half the size of the catches in *High* for every pair of actions chosen. Each entry of the two matrices consists of a vector having an ordered pair on top and a probability vector at the bottom. The ordered pair denotes the payoffs to the players if the corresponding action pair is chosen, the first (second) number is the payoff to player A (B). The probability vector signifies that the system moves to the respective states for the next stage, where the first (second) probability is connected to *High* (*Low*). So, if the play is in state *High* and the players choose action pair $(1, 2)$, the payoff is $\theta_2 = \left(\theta_2^A, \theta_2^B \right)$, i.e., player A receives θ_2^A, player B gets θ_2^B and the play continues next according to $p'_H = (p_H, 1 - p_H)$, i.e., in *High* (respectively *Low*) with probability p_H (respectively $1 - p_H$).

In the sequel, we assume that in both states Action 1 is dominated by the alternative. This places the following restrictions on parameters above:

$$
\begin{aligned}
\theta_1^A &\leq \theta_3^A, \theta_2^A \leq \theta_4^A, \theta_5^A \leq \theta_7^A, \theta_6^A \leq \theta_8^A, \\
\theta_1^B &\leq \theta_2^B, \theta_3^B \leq \theta_4^B, \theta_5^B \leq \theta_6^B, \theta_7^B \leq \theta_8^B.
\end{aligned}
$$

Here, the action pair chosen is not of any influence on the transition probabilities. Moreover, the latter are identical in a given state. Assuming $p_L, p_H \in (0, 1)$, which is sufficient to guarantee that neither of the states is absorbing, we may compute the long-run probabilities of the play visiting *High* respectively *Low* as

$$p = \frac{p_L}{1 - p_H + p_L} \text{ respectively } (1 - p) = \frac{1 - p_H}{1 - p_H + p_L}.$$

Clearly, $p_H = p_L$ implies $p_L = p$. Since the transition probabilities from one state to another for each stage, as well as the long run probability that the play will occur in each state are completely out of the control by both players, we call these variations exogenous.

As the transition probabilities do not depend on the actual choices of the players at any stage, we can regard the game as an infinite sequence of one-shot games where nature decides whether the stage game is

$$\begin{bmatrix} \theta_1 & \theta_2 \\ \theta_3 & \theta_4 \end{bmatrix} \text{ or } \begin{bmatrix} \theta_5 & \theta_6 \\ \theta_7 & \theta_8 \end{bmatrix}.$$

Both one-shot games possess a pure strict Nash equilibrium $(2, 2)$, an infinite repetition of this action pair at all stages yields long-run average payoffs 'converging to' $p\theta_4 + (1 - p)\theta_8$.

2.2 Semi-endogenous variability

To prepare for the sequel, we now assume that the transition probabilities do depend in detail on the actual choices made by the players at that particular stage and state. There exist eight different action pairs which may be chosen, hence the above translates into

$$H = \begin{bmatrix} \begin{pmatrix} \theta_1 \\ p'_1 \end{pmatrix} & \begin{pmatrix} \theta_2 \\ p'_2 \end{pmatrix} \\ \begin{pmatrix} \theta_3 \\ p'_3 \end{pmatrix} & \begin{pmatrix} \theta_4 \\ p'_4 \end{pmatrix} \end{bmatrix} \text{ and } L = \begin{bmatrix} \begin{pmatrix} \theta_5 \\ p'_5 \end{pmatrix} & \begin{pmatrix} \theta_6 \\ p'_6 \end{pmatrix} \\ \begin{pmatrix} \theta_7 \\ p'_7 \end{pmatrix} & \begin{pmatrix} \theta_8 \\ p'_8 \end{pmatrix} \end{bmatrix}$$

So, if the agents choose action pair $(1, 2)$ in *High*, they receive stage payoffs θ_2 and $p'_2 = (p_2, 1 - p_2)$, i.e., the play continues in *Low* with probability $1 - p_2$ next. For Small Fish Wars the following restrictions seem meaningful

$$1 > p_1 \geq p_2 = p_3 \geq p_4$$
$$p_5 \geq p_6 = p_7 \geq p_8 > 0$$
$$p_i \geq p_{i+4} \text{ for } i = 1, ..., 4.$$

We assume that two-sided full restraint causes less damage in both states to the resource than if one player catches with restraint and his opponent

catches without. Hence, the probability that during the next stage play is in *High* if the first case arises is larger than the corresponding probability in the second case. This motivates $p_1 \geq p_2, p_3$ and $p_5 \geq p_6, p_7$. We also assume that it does not matter which player catches with or without restraint, hence the equalities $p_2 = p_3$ and $p_6 = p_7$. Furthermore, we assume that one player catching without restraint is less harmful to the resource than two players catching without restraint. Finally, the inequalities $p_i \geq p_{i+4}$ for $i = 1, ..., 4$, hold because if the play is in *Low*, the system is assumed more vulnerable to overfishing than if the play is in *High*. Catching without restraint by both players in *Low* for instance may make a transition to *High* less likely than if the players were to choose the same action pair in *High*. We refer to e.g., Kelly *et al.* [2006] for an empirical underpinning of these modeling choices. Note that this framework is more general than the former, because the inequalities are chosen such that the former is a special case of the latter.

Let us capture the past play by two matrices

$$QH^t = \left[\begin{array}{cc} q_1^t & q_2^t \\ q_3^t & q_4^t \end{array} \right] \text{ and } QL^t = \left[\begin{array}{cc} q_5^t & q_6^t \\ q_7^t & q_8^t \end{array} \right].$$

Here, e.g., q_1^t is the relative frequency with which action pair top-left in *High* has occurred during past play, and q_7^t is the relative frequency of action pair bottom-left in *Low* having occurred during past play. Given this interpretation we must have

$$q^t = \left(q_1^t, ..., q_8^t \right) \in S^7, \text{ for all } t = 1, 2, ...,$$

where S^n denotes an n-dimensional unit simplex.

Assume that

$$\lim_{t \to \infty} QH^t = \left[\begin{array}{cc} q_1 & q_2 \\ q_3 & q_4 \end{array} \right] \text{ and } \lim_{t \to \infty} QL^t = \left[\begin{array}{cc} q_5 & q_6 \\ q_7 & q_8 \end{array} \right],$$

Then it must hold also that $\sum_{i=1}^{8} q_i = 1$. Furthermore, play jumped from *High* to *Low* with a relative frequency of $\sum_{i=1}^{4} q_i (1 - p_i)$ and from *Low* to *High* with a relative frequency of $\sum_{i=5}^{8} q_i p_i$. If at least one of the transition probabilities $p_1, ..., p_4$ is strictly smaller than one and at least one of the transition probabilities $p_5, ..., p_8$ is strictly larger than zero, an additional restriction on the system must hold, namely

$$\sum_{i=1}^{4} q_i (1 - p_i) = \sum_{i=5}^{8} q_i p_i. \tag{1}$$

If these matrices QH^t and QL^t converge for $t \to \infty$, we may regard $q = (q_1, ..., q_8)$ as a stationary distribution over the choices of all action pairs in both states. Not all such 'stationary distributions' are admissible as (1) must be fulfilled.

6

Example 1a First, we connect actual numbers to the pairs of stage payoffs received by the players.

$$H = \begin{bmatrix} \begin{pmatrix} (4,4) \\ (p_1, 1-p_1) \end{pmatrix} & \begin{pmatrix} \left(\frac{7}{2}, 6\right) \\ (p_2, 1-p_2) \end{pmatrix} \\ \begin{pmatrix} \left(6, \frac{7}{2}\right) \\ (p_3, 1-p_3) \end{pmatrix} & \begin{pmatrix} \left(\frac{11}{2}, \frac{11}{2}\right) \\ (p_4, 1-p_4) \end{pmatrix} \end{bmatrix},$$

$$L = \begin{bmatrix} \begin{pmatrix} (2,2) \\ (p_5, 1-p_5) \end{pmatrix} & \begin{pmatrix} \left(\frac{7}{4}, 3\right) \\ (p_6, 1-p_6) \end{pmatrix} \\ \begin{pmatrix} 3, \frac{7}{4} \\ (p_7, 1-p_7) \end{pmatrix} & \begin{pmatrix} \frac{11}{4}, \frac{11}{4} \\ (p_8, 1-p_8) \end{pmatrix} \end{bmatrix}.$$

We now compare two strategy pairs. In the first strategy pair, both players use restraint at all stages of the play, and in the second one both players use no-restraint at all stages of the play. Then, under the first pair of strategies the restriction above implies

$$q_1(1 - p_1) = q_5 p_5.$$

from which we derive $q_1 = \frac{p_5}{1-p_1+p_5}$. By analogy we find for the second case that $q_4 = \frac{p_8}{1-p_4+p_8}$. Under the assumptions made, q_1 is precisely the long-run probability that *High* occurs under perfect restraint, and q_4 is the long-run probability that *Low* occurs under total absence of restraint. It can be confirmed under the restrictions imposed that $q_1 \geq q_4$ since

$$\begin{aligned} q_1 - q_4 &= \frac{p_5}{1-p_1+p_5} - \frac{p_8}{1-p_4+p_8} \\ &= \frac{p_5(1-p_4+p_8) - p_8(1-p_1+p_5)}{(1-p_1+p_5)(1-p_4+p_8)} \\ &= \frac{p_5(1-p_4) - p_8(1-p_1)}{(1-p_1+p_5)(1-p_4+p_8)}. \end{aligned}$$

As $p_5 \geq p_8$ and $(1-p_4) \geq (1-p_1)$, this boils down to $p_5(1-p_4) - p_8(1-p_1) \geq 0$, hence the inequality $q_1 \geq q_4$ is satisfied. ∎

Observe that all transition probabilities are given exogenously, yet the long-run probability that *High* (*Low*) occurs during the play depends crucially on the strategy pair employed by the players. This is why we used the term semi-endogenous variations in the title of this subsection. So, in the case of full restraint in the example, the system spends more time in *High* than in the case of no-restraint.

Example 1b To continue, let $p_1 = 0.8$, $p_2 = p_3 = 0.6$, $p_4 = 0.3$, $p_5 = 0.6$, $p_6 = p_7 = 0.4$ and $p_8 = 0.1$. We find then

$$q_1 = \frac{0.6}{1-0.8+0.6} = \frac{3}{4}, \ q_4 = \frac{0.1}{1-0.3+0.1} = \frac{1}{8}.$$

So, under perfect restraint the play is six times more often in *High* than under no-restraint. An interesting phenomenon now occurs if one compares the long-term average payoffs connected to these two qualitatively opposed strategy pairs. These averages are of course given by:

$$q_1(4,4) + q_5(2,2) = \frac{3}{4}(4,4) + \frac{1}{4}(2,2) = \left(3\frac{1}{2}, 3\frac{1}{2}\right),$$

$$q_4\left(\frac{11}{2}, \frac{11}{2}\right) + q_8\left(\frac{11}{4}, \frac{11}{4}\right) = \frac{1}{8}\left(\frac{11}{2}, \frac{11}{2}\right) + \frac{7}{8}\left(\frac{11}{4}, \frac{11}{4}\right) = \left(3\frac{3}{32}, 3\frac{3}{32}\right).$$

So, the strategy pair giving lower payoffs to both players in every stage game actually induces higher average payoffs in the long run. ∎

This example may serve to demonstrate that we already have the makings of a social dilemma of the social trap variety (cf., e.g., Cross & Guyer [1980], see Komorita & Parks [1994] for an overview). The long-term average payoffs associated with full restraint are higher than those connected to no-restraint. So, despite the circumstance that no-restraint is the dominant action for both players in each stage game, the continued use of this one-shot-optimal action pair leads to lower average payoffs in the long run than perfect restraint does. Games with similar structures draw wide interest in (political) economics, psychology, sociology and law.[4]

3 Strategies and rewards

The players receive an infinite stream of stage payoffs during the play, and 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

$$\gamma^k(\pi, \sigma) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} R_t^k(\pi, \sigma).$$

Let furthermore, $\gamma(\pi, \sigma) \equiv \left(\gamma^A(\pi, \sigma), \gamma^B(\pi, \sigma)\right)$. First, we focus on rewards from strategies which are pure and jointly convergent. Then, we extend our analysis to obtain larger sets of feasible rewards.

A strategy is **pure**, if at *each* stage a **pure action** is chosen, i.e., an action is chosen with probability 1. The set of pure strategies for player k is $\mathcal{P}^k$, and $\mathcal{P} \equiv \mathcal{P}^A \times \mathcal{P}^B$. Let us define the following notions, introduced before in a rather informal manner, a bit more formally. For $j = 1, 2$, define

$$q_j^t = \frac{\#\left\{\left(j_u^{A,H}, j_u^{B,H}\right) = (1, j) \mid 1 \leq u \leq t\right\}}{t},$$

---

[4]Cf., e.g., Olson [1965], Hamburger [1973], Platt [1973], Schelling [1978], Dawes [1980], Liebrand [1983], Ostrom [1990], Ostrom *et al.* [1994] or Skyrms [2004]

$$q_{j+2}^t = \frac{\#\left\{\left(j_u^{A,H}, j_u^{B,H}\right) = (2,j) \mid 1 \le u \le t\right\}}{t},$$

$$q_{j+4}^t = \frac{\#\left\{\left(j_u^{A,L}, j_u^{B,L}\right) = (1,j) \mid 1 \le u \le t\right\}}{t},$$

$$q_{j+6}^t = \frac{\#\left\{\left(j_u^{A,H}, j_u^{B,H}\right) = (2,j) \mid 1 \le u \le t\right\}}{t}.$$

Here, $j_u^{A,X}$ $(j_u^{B,X})$ denotes the action taken by player A (B) while being in state $X = H, L$ at stage $u \in \{1, 2, ..., t\}$. So, for instance q_4^t is the relative frequency of action pair $(2,2)$ in state H being chosen until stage t. Hence,

$$q^t \equiv \left(q_1^t, ..., q_8^t\right) \in S^7 = \left\{x \in \mathbb{R}_+^{n+1} \mid \sum_{i=1}^{8} x_i = 1\right\}.$$

We refer to such a vector as the **relative frequency vector**. Moreover, for vector $q \in S^7$, the q-**averaged payoffs** $(x,y)_q$ are given by

$$(x,y)_q = \sum_{i=1}^{8} q_i \theta_i.$$

The strategy pair $(\pi, \sigma) \in \mathcal{X}^A \times \mathcal{X}^B$ is **jointly convergent** if and only if $q \in S^7$ exists such that for all $\varepsilon > 0$, $i \in \{1, 2, ..., 8\}$:

$$\limsup_{t \to \infty} \Pr_{\pi,\sigma}\left[\left|q_i^t - q_i\right| \ge \varepsilon\right] = 0. \tag{2}$$

$\Pr_{\pi,\sigma}$ denotes the probability under strategy-pair (π, σ). $\mathcal{JC}$ denotes the set of jointly-convergent strategy pairs. Under such a pair of strategies, the relative frequency of each action pair in both states as play goes to infinity converges to a fixed number with probability 1 in the terminology of Billingsley [1986, p.274]). The **set of jointly-convergent pure-strategy rewards** $P^{\mathcal{JC}}$ is then the set of pairs of rewards obtained by using a pair of jointly-convergent strategies.

The following result connects the notions introduced.

Proposition 1 *Let strategy-pair* $(\pi, \sigma) \in P^{\mathcal{JC}}$ *and let* $q \in S^7$ *for which (2) is satisfied, then the average payoffs are given by* $\gamma(\pi, \sigma) = (x,y)_q$.

Proof: Let $(\pi, \sigma) \in P^{\mathcal{JC}}$ and denote the pair of expected payoffs received at stage u under this strategy pair by $E\{\theta_u^{\pi,\sigma}\}$, then

$$\lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^{t} E\{\theta_u^{\pi,\sigma}\} = \lim_{t \to \infty} E\left\{\frac{1}{t} \sum_{u=1}^{t} \theta_u^{\pi,\sigma}\right\} =$$

$$\lim_{t \to \infty} E \left\{ \sum_{i=1}^{8} q_i^t \theta_i \right\} = \lim_{t \to \infty} \sum_{i=1}^{8} E \left\{ q_i^t \right\} \theta_i = \sum_{i=1}^{8} q_i \theta_i = (x, y)_q.$$

The second equality sign holds because it involves a change in counting: on the left-hand side we sum over all periods, on the right-hand side over all eight entries of the two bi-matrices weighed by their relative frequencies. The first and third equality are standard, the penultimate one follows from (2), cf., e.g., Billingsley [1986, p.274], the final one by the definition given above. Since $\lim_{t \to \infty} \frac{1}{t} \sum_{u=1}^{t} E\{\theta_u^{\pi,\sigma}\}$ equals $(x, y)_q$, it follows that $\gamma(\pi, \sigma) = (x, y)_q$. ∎

We now show implications of this result for two variants of a Small Fish War, namely one with exogenous and one with semi-endogenous stochasticity.
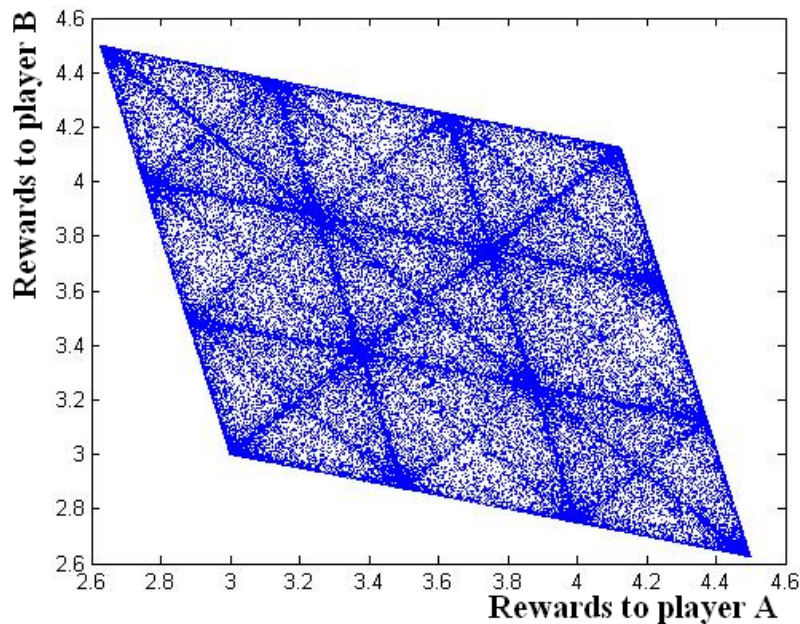

Figure 1: A sketch of all jointly-convergent pure-strategy rewards in Example 1a with $p_1 = ... = p_8 = 0.5$. Matlab generated 90.000 of such rewards randomly. The true set is dense.

Example 2 Figure 1 shows $P^{\mathcal{JC}}$, the set of jointly-convergent pure-strategy rewards, if, as in the case of Subsection 2.1, the transition probabilities are exogenously given by $p_H = p_L = 0.5$. Please recall that in the long run the system spends half of the time in *High*. Since the payoffs in *Low* were originally selected as being 50% of those in *High*, it may be seen readily

that these rewards coincide with rewards being 75% ($= \frac{1}{2} \cdot 100\% + \frac{1}{2} \cdot 50\%$) of those if the game had only one state, namely High. This rather intuitive observation is far from generalizable to games with different choices of transition probabilities, as we show next.
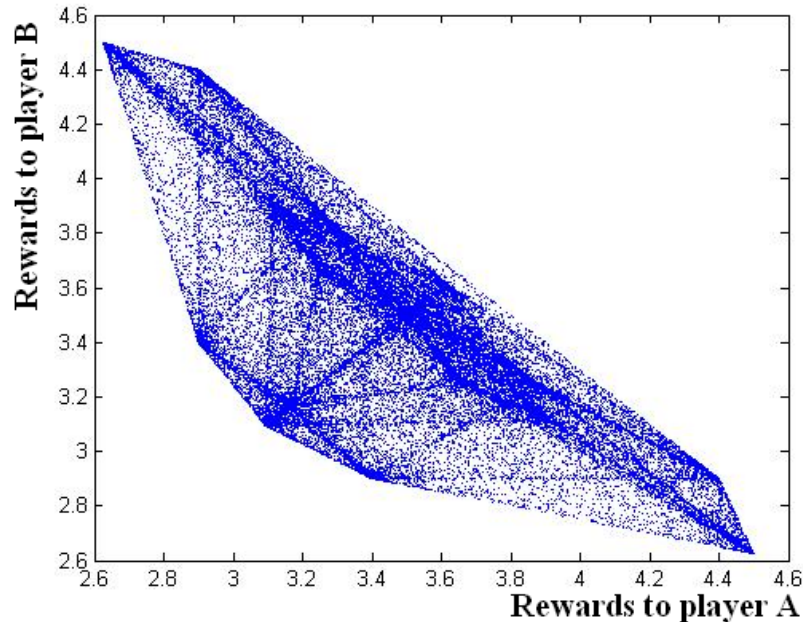

Figure 2: A sketch of $P^{\mathcal{JC}}$ resembles a projection of a diamond for the case $p_1 = 0.8$, $p_2 = p_3 = 0.6$, $p_4 = 0.3$, $p_5 = 0.6$, $p_6 = p_7 = 0.4$, $p_8 = 0.1$. Matlab generated 50,000 of such rewards randomly. The true set is dense.

Figure 2 shows $P^{\mathcal{JC}}$ for a case labeled as semi-endogenous variations in Subsection 2.2. The transition probabilities for this numerical example are given in the caption of the figure. As we have argued, the time the play occurs in *High* depends crucially on the strategies the players employ, despite the transition probabilities themselves do not change. Therefore, also the rewards depend on those long run probabilities the play occurs in either state.

For Figures 1 and 2 only the transition probabilities differ. The quadrangle of Figure 1 is distorted into a diamond shape by these different transition probabilities. Note that the strategy pair consisting of always catching with restraint yields $\left(3\frac{1}{2}, 3\frac{1}{2}\right)$, whereas no restraint yields $\left(3\frac{3}{32}, 3\frac{3}{32}\right)$ as shown before. So, the quadrangle is folded inwards, so to speak, from the north east as well as from the south west, such that the utmost south-western point in the original ends up more to the north-east than the original utmost north-eastern point. ∎

4 Endogenous transition probabilities

Central for the approach followed in this section is that we endogenize the transition probabilities in the sense that we allow them to depend on past play. To be a bit more precise, we take the transition probabilities as follows

$$p_i^t = f_i \left(q^t \right) \ \text{ for all } i = 1, ..., 8.$$

So, for instance, p_7^t is the probability that the play moves to *High* at stage $t + 1$ given that the players played action pair $(2, 1)$ in *Low* at point t in time given the history of the play captured by relative frequency vector q^t. The functions $f_i : S^n \to [0, 1]$ are assumed continuous for all $i = 1, ..., 8$.

Our motivation for this innovation stems from a desire to incorporate empirically observed phenomena into Small Fish Wars. One striking observation made in real-life renewable resource systems is that the resource can be brought down in numbers (or quantity) rather quickly by overexploitation. Yet, when more sound catching policies (strategies) are adopted designed at restoring the resource, it takes quite long for the system to recover to higher or full capacity.[5] For this purpose, we design the new game in such a matter that *Low*, which is not an 'absorbing' state per se, may become one if both agents use no-restraint for prolonged periods.

By absorbing state in this context, we mean that the system can not move out of the state *temporarily* as opposed to more standard approaches where absorbing means that the system can *never* escape. Now, the state *Low* will be absorbing (in our sense) at stage $t \in \mathbb{N}$ if

$$f_i \left(q^t \right) = 0 \text{ for all } i = 5, ..., 8.$$

So, the values of all of the functions corresponding to action choices in *Low* above must be zero at that point in time. Of course, if one of these values becomes strictly positive again, the play may escape from this state if the agents target the corresponding action pair. Assuming relative frequency vectors exist for which *Low* is not absorbing, the agents must manipulate the current relative frequency vector sufficiently close towards them. This may take some time, and meanwhile only bad payoffs occur.

We assume the following ranking to hold for all $t \in \mathbb{N}$:

$$1 > p_1^t \geq p_2^t = p_3^t \geq p_4^t$$
$$p_5^t \geq p_6^t = p_7^t \geq p_8^t \geq 0$$
$$p_i^t \geq p_{i+4}^t \text{ for } i = 1, ..., 4.$$

Note that we have introduced one small but important difference to the setup used in the previous sections, as we do not exclude anymore that *Low*

---

[5]Cf., e.g., Tegner [1992], Hutchings [2000], Russ & Alcala [1996, 2003], Bulte [2003], Kelly *et al.* [2006], Courchamp *et al.* [2006], Hall *et al.* [2008]). Another well-known low-density phenomenon, the Allee Effect, was modeled in Joosten [2007b,c, 2010].

becomes an absorbing state (temporarily). Otherwise, we assume the same ordinal ranking of the transition probabilities to hold as before, but now, since they can change, for any $t \in \mathbb{N}$.

Example 3a Let for given relative frequency vector $q^t \in S^7$, the transition functions f_i, $i = 1, ..., 8$, governing the transition probabilities, be given by

$$
\begin{array}{l}
p_1^t = f_1(q^t) = \left[\frac{1}{2} - \frac{11}{24}q_4^t - \frac{11}{12}q_8^t\right]_+ \\
p_2^t = f_2(q^t) = p_3^t = f_3(q^t) = \left[\frac{1}{2} - \frac{11}{20}q_4^t - \frac{11}{10}q_8^t\right]_+ \\
p_4^t = f_4(q^t) = \left[\frac{1}{2} - \frac{11}{16}q_4^t - \frac{11}{8}q_8^t\right]_+ \\
p_5^t = f_5(q^t) = \left[\frac{1}{2} - \frac{11}{12}q_4^t - \frac{11}{6}q_8^t\right]_+ \\
p_6^t = f_6(q^t) = p_7^t = f_7(q^t) = \left[\frac{1}{2} - \frac{11}{8}q_4^t - \frac{11}{4}q_8^t\right]_+ \\
p_8^t = f_8(q^t) = \left[\frac{1}{2} - \frac{11}{4}q_4^t - \frac{11}{2}q_8^t\right]_+ .
\end{array}
$$

Here, $[x]_+$ is short-hand for $\max\{x, 0\}$. These equations capture the following ideas. If the fishermen were never to interfere in the fish stock, the probability that the play at any stage is in *Low* (*High*) equals $\frac{1}{2}$. However, the catching behavior of the agents may have consequences for the probabilities of both states occurring during the play. Only unrestrained catching by both agents damages the resource. In *Low*, the effects of combined unrestrained catching are more detrimental to the transition probabilities than in *High*.

Suppose both agents play action 1 twice followed by action 2 for a sufficiently long period of time. Clearly, $q_4^t + q_8^t = \frac{t-2}{t}$, hence for t large enough $f_5(q^t) = f_6(q^t) = f_7(q^t) = f_8(q^t) = 0$, because

$$
\frac{1}{2} - \frac{11}{12}q_4^t - \frac{11}{6}q_8^t = \frac{1}{2} - \frac{11}{12}\left(\frac{t-2}{t} - q_8^t\right) - \frac{11}{6}q_8^t =
$$

$$
\frac{1}{2} - \frac{11}{12}\left(1 - \frac{2}{t} - q_8^t\right) - \frac{11}{6}q_8^t = -\frac{5}{12} + \frac{11}{6t} - \frac{11}{12}q_8^t < 0.
$$

Then, $f_5(q^t) = 0$ and by the relation to the other transition probability functions, $f_6(q^t) = f_7(q^t) = f_8(q^t) = 0$ as well. Take $t^* = 5$, clearly

$$
-\frac{5}{12} + \frac{11}{6t^*} - \frac{11}{12}q_8^{t^*} < 0,
$$

because even if the state *Low* did not materialize in the first four periods, hence $q_8^{t^*} = 0$, the part $\frac{11}{6t^*}$ is sufficiently small.

It should be noted that even if both agents switch to playing a pair of sequences of $(1, 1, 1, ...)$ at some point $t^{**} > t^*$, it will take a while before $f_5(q^t)$ becomes positive again.[6] If the agents were to switch to a seemingly equally 'harmless' pair of sequences alternating their actions in an anti-coordinating

---

[6]Shertzer & Prager [2006] point out that in real life delays regarding switching to stock rebuilding measures in fishery management are the rule rather than the exception.

fashion, the play would eventually return to *High* again, but after a longer period of time (in expectation) than under the pair of sequences $(1, 1, 1, ...)$ as $f_6\left(q^t\right)$ and $f_7\left(q^t\right)$ become strictly positive at an even later point in time. While being in *Low* they receive a stream of stage payoffs well below the lowest stage payoff in *High*.   ■

This example shows that one of the aspects of real-life resource systems can be replicated to some degree, namely renewable resources recuperating slowly after a program of recovery has been taken up. The reason for this may be in biological specifics. For instance, it is well-known (cf., e.g., Kurlansky [1998]) that full-grown cod spawn a considerably higher number of eggs than younger specimen. Oosthuizen & Daan [1974] find that linear fecundity-weight relations have highest statistical explanatory power. Armstrong *et al.* [2001] confirm these findings and demonstrate that length has rather poor correlation with fecundity. Rose *et al.* [2008] claim that fecundity increases with size in an exponential manner. Despite these differences, these studies agree on 'big is beautiful' when it comes to reproductive capacity. As mature, big fish are targeted and modern techniques such as gill nets facilitate this, the effects of overfishing are mainly felt in the cohorts most productive in providing offspring. For a species to regain full reproductive capacity, younger cohorts must reach ages well-beyond adulthood. This may take a considerable while.

For a pair of jointly-convergent pure strategies, i.e., a pair of strategies such that (1) holds, we clearly have

$$p_i \equiv \lim_{t\to\infty} p_i^t = \lim_{t\to\infty} f_i\left(q^t\right) = f_i(q) \text{ for } i = 1, ..., 8.$$

Again this can be motivated by these functions being continuous, and the strong property formulated in (1) (cf. e.g., Billingsley [1986]). The notation introduced above generalizes the one used in the previous sections.

To establish $P^{\mathcal{JC}}$ we distinguish the following three possibilities

$$\sum_{i=1}^{4} q_i\left(1 - p_i\right) = \sum_{i=5}^{8} q_i p_i,$$

$$\text{or } \sum_{i=5}^{8} q_i p_i = 0,$$

$$\text{or } \sum_{i=1}^{4} q_i\left(1 - p_i\right) = 0.$$

The first equation pertains to neither state being absorbing in the long run, the second one to *Low* being an absorbing state in the long run, and the final one to *High* being absorbing. Observe that the three possibilities are

14

mutually exclusive, since for $q \in S^7$, the underlying restriction is $\sum_{i=1}^{8} q_i = 1$. Now, the second equality can only hold if

$$p_i = 0 \text{ or } q_i = 0 \text{ for all } i = 5, ..., 8.$$

Similarly, the final situation can only hold if

$$1 - p_i = 0 \text{ or } q_i = 0 \text{ for all } i = 1, ..., 4.$$

So, *if a state is absorbing, then positive mass on a component of the relative frequency vector q can only occur if the associated probability of leaving that state is zero.*

Note that *High* can not be an absorbing state, but *Low* may very well become one during the play. Clearly, if *Low* is absorbing it must hold that

$$\sum_{i=5}^{8} q_i = 1.$$

Taking this fact and the ranking assumed together, we may distinguish the following three subcases if *Low* is an absorbing state.

$$q_8 = 1 \text{ and } p_6 = p_7 > p_8 = 0,$$

$$\text{or } \sum_{i=6}^{8} q_i = 1 \text{ and } p_6 = p_7 = p_8 = 0,$$

$$\text{or } \sum_{i=5}^{8} q_i = 1 \text{ and } p_5 = p_6 = p_7 = p_8 = 0.$$

We will now demonstrate how the effects mentioned may be replicated by a stochastic game model with endogenous transition probabilities. Combined unrestrained catching in *Low* is more damaging to the transition probabilities to *High* than the same behavior does in *High*.

Example 3b The set of jointly-convergent pure-strategy rewards depends crucially on whether *Low* is an absorbing state or not. Let (π, σ) be a pair of jointly-convergent pure strategies and let q satisfy (2). Then, the equation

$$\sum_{i=1}^{4} q_i \left(1 - f_i(q)\right) = \sum_{i=5}^{8} q_i f_i(q) \text{ and } f_8(q) > 0$$

places easy to check restrictions on the system for the case that *Low* is not absorbing. For the case that it is, we must distinguish three subcases. *Low* is an absorbing state iff one of the following three conditions are fulfilled:

$$q_8 = 1 \text{ and } \frac{1}{2} - \frac{11}{4}q_4 - \frac{11}{2}q_8 \leq 0 < \frac{1}{2} - \frac{11}{8}q_4^t - \frac{11}{4}q_8^t,$$

$$\text{or } \sum_{i=6}^{8} q_i = 1 \text{ and } \frac{1}{2} - \frac{11}{8}q_4 - \frac{11}{4}q_8 \leq 0 < \frac{1}{2} - \frac{11}{12}q_4^t - \frac{11}{6}q_8^t,$$

15

$$\text{or } \sum_{i=5}^{8} q_i = 1 \text{ and } \frac{1}{2} - \frac{11}{12}q_4 - \frac{11}{6}q_8 \leq 0.$$

Observe that $q_4 > 0$ yields a contradiction. Hence, the above reduces to

$$q_8 = 1 \text{ and } \frac{1}{11} \leq q_8 \leq \frac{2}{11},$$

$$\text{or } \sum_{i=6}^{8} q_i = 1 \text{ and } \frac{2}{11} \leq q_8 \leq \frac{3}{11},$$

$$\text{or } \sum_{i=5}^{8} q_i = 1 \text{ and } q_8 \geq \frac{3}{11}.$$

Clearly, the first of three cases yields a contradiction, so there exists actually not a single relative frequency vector $q \in S^7$ leading to the situation that *Low* is an absorbing state such that **only** $f_8(q) = 0$. The second implies that $q_5 = 0$ and $\frac{1}{24} \leq q_8 \leq \frac{1}{20}$. For $q_8 \geq \frac{1}{20}$, all other relative frequencies $q_5, ..., q_7$ may be non-zero as well and simultaneously fulfill the criteria under which *Low* is an absorbing state. In Figure 3 we have depicted the sets of admissible relative frequency vectors consistent with the case that the state *Low* is an absorbing state.
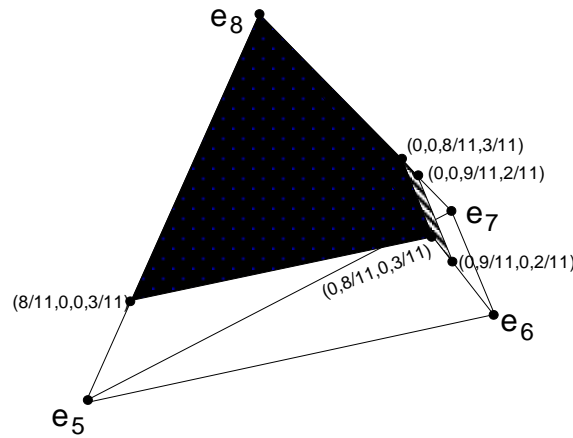

Figure 3: If *Low* is absorbing, all mass is on $q_5, ..., q_8$. We depict S^3. a facet of S^7; each point in S^3 corresponds to an admissible vector q. Extreme point e_i has element i equal to one. There are two sets of admissible q's, sketched here as a 3-dimensional set, and a 2-dimensional boundary set.

Obviously, such a reduction of possible relative frequency vectors has consequences on the rewards which can be obtained under the assumption

that the strategies involved make *Low* an absorbing state. Figure 4 depicts all jointly-convergent pure-strategy rewards if *Low* is an absorbing state and the relative frequency vectors are admissible. Note that the set of rewards obtained is **not convex**.
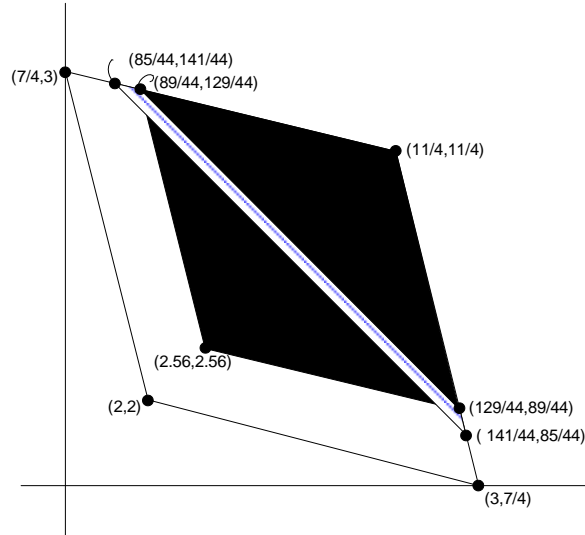

Figure 4: The parallellogram corresponds to 3-dimensional set, the trapezium to the 2-dimensional boundary set of admissible q's.

A second component of the set of jointly-convergent pure-strategy rewards is also determined quite easily, namely those consistent with strategies where mutual no-restraint is avoided. Since this leads to $q_4 = q_8 = 0$, it follows that with respect to the transition probabilities, we have:

$$f_i(q) = \frac{1}{2} \text{ for all } i = 1, ..., 8.$$

It can be confirmed that the set of jointly-convergent pure-strategy rewards corresponding to relative frequency vector with $q_4 = q_8 = 0$ is given by

$$conv \left\{ (3,3), \left(\frac{18}{4}, \frac{21}{8} \right), \left(\frac{21}{8}, \frac{18}{4} \right) \right\}.$$

Here, *conv S* denotes the convex hull of set S.

The remaining jointly-convergent pure-strategy rewards are depicted in Figure 5. The figure was generated in Matlab. As the program only generates a large but finite number of points representing jointly-convergent pure-strategy rewards, the set does not appear dense at the present settings. The true set is dense and we could make it look as such by adjusting the parameters. However, presently the two sets described show up clearly,
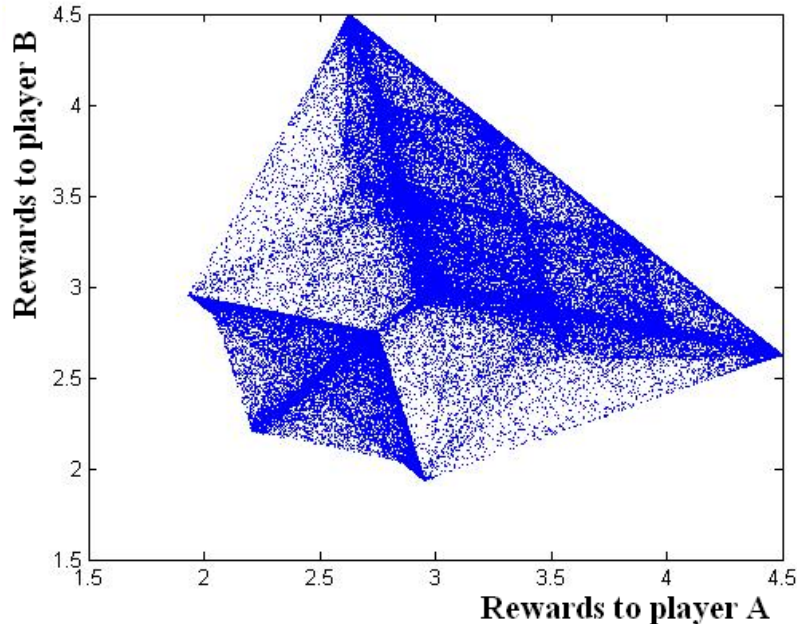
Figure 5: The set $P^{\mathcal{JC}}$ for Examples 3a and 3b. The hexagon and the triangle discussed, can be distinguished clearly. The area in between corresponds to rewards obtained by strategy-pairs with positive weight on the two action-pairs damaging the resource.

an effect lost if the figure were to be made more dense. It turns out that the space in between the two subsets elaborated on, is filled by the remaining jointly-convergent pure-strategy rewards. ∎

5   On equilibrium rewards

We will now focus on rewards which may be obtained by equilibria involving threats. The approach to be followed is similar to a well-established one in the repeated games literature (cf., e.g., Hart [1985], Forges [1986]), where it is linked to the Folk Theorem (see e.g., Van Damme [1992]). Applications of this approach to stochastic games exist as well (cf., e.g., Thuijsman & Vrieze [1998], Joosten *et al.* [2003], Schoenmakers [2004], Joosten [1996, 2005, 2007a,b]), but they have attracted much less following than in the branch of literature mentioned.

We call $v = \left(v^A, v^B\right)$ 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 A's average payoffs. Under a pair of **individually rational** (feasible) rewards each player receives at least the

18

threat-point reward. We can now present the following formal result.

Theorem 2 *Let E be the set of all individually-rational jointly-convergent pure-strategy rewards. Then, each pair of rewards in E can be supported by an equilibrium.*

Proof: Let π^p (σ^p) be a punishment-strategy of player A (B), i.e., a strategy holding his opponent to at most v^2 (v^1). Let $(x, y) \in E$, then $(\pi, \sigma) \in P^{\mathcal{JC}}$ exists such that $\gamma(\pi, \sigma) = (x, y) \geq (v^1, v^2)$. Define (π^*, σ^*) by

$$\pi_t^* = \begin{cases} \pi_t & \text{if } j_k = \sigma_k^* \text{ for all } k < t, \\ \pi_t^p & \text{otherwise.} \end{cases}$$

$$\sigma_t^* = \begin{cases} \sigma_t & \text{if } i_k = \pi_k^* \text{ for all } k < t, \\ \sigma_t^p & \text{otherwise.} \end{cases}$$

Here, i_t and j_t denote the actions taken by player A respectively B at stage t of the play. Clearly, $\gamma(\pi^*, \sigma^*) = \gamma(\pi, \sigma) = (x, y)$. Suppose Player 1 were to play π' such that $\pi_k' \neq \pi_k^*$ for some k, then Player 2 would play according to σ^p from then on. Since, $\gamma^1(\pi', \sigma^p) \leq v^1 \leq x$, it follows immediately that player A can not improve against σ^*. A similar statement holds in case Player 2 deviates unilaterally. Hence, (π^*, σ^*) is an equilibrium. ∎

The above is fairly standard in the analysis of repeated games, see e.g., Hart [1985], Van Damme [1991], and elsewhere, e.g., Thuijsman & Vrieze [1998]. Such a pair of strategies (π^*, σ^*) is commonly referred to as an equilibrium involving threats. The threat is that each player will punish the opponent if the latter were to deviate from the equilibrium path.

In Joosten *et al.* [2003], we presented a technique to prove that the convex hull of a set of E correspond to equilibria as well. The convexification is based on a communication round deciding on which of the equilibria yielding a reward belonging to E is to be played from then on, see also Hart [1985] and Forges [1986] on the issue of obtaining (convex) sets of equilibrium rewards. Using that technique it can be shown that the convex hull of E is contained by the set of equilibrium rewards. With some mild genericity assumptions it can be shown that except for Pareto-inferior boundaries of this convex hull, this set can be supported by so-called subgame perfect equilibria, cf., e.g., Joosten *et al.* [2003].

Next, we illustrate the notions introduced and Theorem 2.

Example 4a In the case underlying Figure 1, the threat point is given by $v = (v^A, v^B) = \left(\frac{33}{8}, \frac{33}{8}\right)$. This can be seen as follows. Suppose player A plays the second action from a certain stage onwards. Then, player B faces a rather simple problem from then on, he must choose between obtaining $\frac{7}{2}$ and $\frac{11}{2}$ in *High,* and between $\frac{7}{4}$ and $\frac{11}{4}$ in *Low* where each state occurs with

probability $\frac{1}{2}$. So, B maximizes his reward by always choosing the second action and the rewards must equal

$$\frac{1}{2} \cdot \frac{11}{2} + \frac{1}{2} \cdot \frac{11}{4} = \frac{33}{8}.$$

So, $v^B \leq \frac{33}{8}$. Note furthermore that if player B were to play his second action at every stage of the play, then player A can minimize his opponent's reward by choosing this action as well. Hence, player B has a strategy to guarantee himself at least $\frac{33}{8}$, so $v^B \geq \frac{33}{8}$. ∎

The analysis not only confirms our earlier statement about no-restraint being an equilibrium, but we showed that it is the unique equilibrium. Existence of multiple equilibria is not an exception as the next example shows.

Example 4b In the case underlying Figure 2, we have $v = \left(v^A, v^B\right) = \left(\frac{137}{44}, \frac{137}{44}\right)$. This can be seen as follows. Suppose player A plays the second action from a certain stage onwards. Then, player B faces the following maximization problem (cf., e.g., Hordijk *et al.* [1983], Blackwell [1962]):

$$\max_{q_3, q_4, q_7, q_8} \frac{7}{2}q_3 + \frac{11}{2}q_4 + \frac{7}{4}q_7 + \frac{11}{4}q_8$$
$$\text{s.t. } 1 = q_3 + q_4 + q_7 + q_8$$
$$0 = \frac{4}{10}q_3 + \frac{7}{10}q_4 - \frac{4}{10}q_7 - \frac{1}{10}q_8$$
$$0 \leq q_3, q_4, q_7, q_8.$$

From the linearity of this problem it follows immediately that it suffices to examine the four situations in which one of two frequencies in *High* is maximal combined with one of two frequencies in *Low* is maximal. The solution is $q_7 = \frac{7}{11}$ and $q_4 = \frac{4}{11}$, and the B's reward is $\frac{4}{11} \cdot \frac{11}{2} + \frac{7}{11} \cdot \frac{7}{4} = \frac{137}{44}$. So, $v^B \leq \frac{137}{44}$.

Suppose player B uses his first action in *Low* and his second action in *High*. To minimize his opponent's reward, Player A faces the following minimization problem (cf., e.g., Hordijk *et al.* [1983], Blackwell [1962]):

$$\min_{q_3, q_4, q_7, q_8} 6q_2 + \frac{11}{2}q_4 + 2q_5 + \frac{7}{4}q_7$$
$$\text{s.t. } 1 = q_2 + q_4 + q_5 + q_7$$
$$0 = \frac{4}{10}q_2 + \frac{7}{10}q_4 - \frac{6}{10}q_5 - \frac{4}{10}q_7$$
$$0 \leq q_2, q_4, q_6, q_8.$$

The solution is $q_7 = \frac{7}{11}$ and $q_4 = \frac{4}{11}$. So, $v^B \geq \frac{137}{44}$. Therefore, player B has a strategy guaranteeing him at least this amount no matter what his opponent does, whereas the latter has a strategy by which he can keep his opponent to at most the same amount. Hence, $v^B = \frac{137}{44}$. ∎
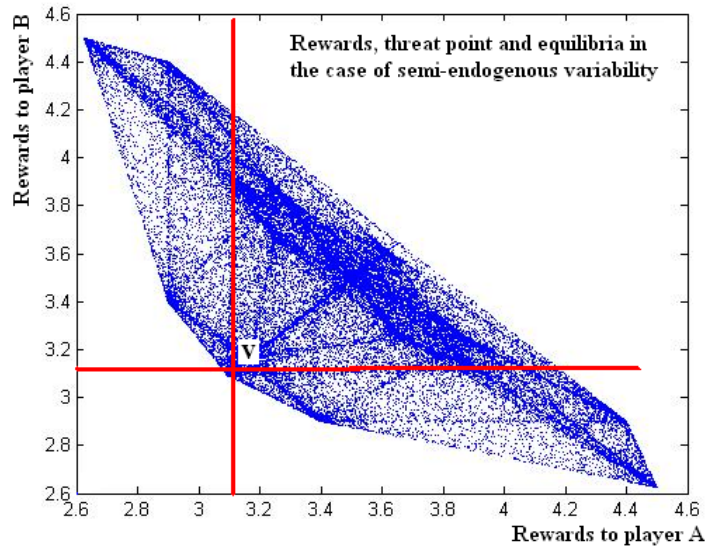
Figure 6: All jointly-convergent pure-strategy rewards located above and to the right of the two red lines through $\left(\frac{137}{44}, \frac{137}{44}\right)$ can be supported by an equilibrium strategy using threats.

The twice-quoted result of Hordijk *et al.* [1983] states that a pure stationary strategy suffices as a best reply against a fixed stationary strategy. Any pair of stationary strategies here is clearly jointly-convergent. For the case in which the transition probabilities are endogenous on past play, the linear programs will not be sufficient as we demonstrate next.

Example 4c. In the case underlying Figure 5, we have $v = \left(v^A, v^B\right) = (2.75, 2.75)$. This can be seen as follows. Suppose player A plays the second action from a certain stage onwards. By the results of Hordijk *et al.* [1983] a pure stationary strategy suffices for player B to maximize his rewards. So, player B faces the following problem

$$
\max_{q_3, q_4, q_7, q_8} \frac{7}{2} q_3 + \frac{11}{2} q_4 + \frac{7}{4} q_7 + \frac{11}{4} q_8
$$

$$
\text{s.t. } 1 \;=\; q_3 + q_4 + q_7 + q_8
$$

$$
0 \;=\; (1 - f_3) q_3 + (1 - f_4) q_4 - f_7 q_7 - f_8 q_8
$$

$$
f_3 \;=\; \left[\frac{1}{2} - \frac{11}{20} q_4 - \frac{11}{10} q_8 \right]_+
$$

$$
f_4 \;=\; \left[\frac{1}{2} - \frac{11}{16} q_4 - \frac{11}{8} q_8 \right]_+
$$

21

$$f_7 = \left[\frac{1}{2} - \frac{11}{8}q_4 - \frac{11}{4}q_8\right]_+$$

$$f_8 = \left[\frac{1}{2} - \frac{11}{4}q_4 - \frac{11}{2}q_8\right]_+$$

$$0 \leq q_3, q_4, q_7, q_8.$$

Despite the linearity of the objective function, we do not have an *LP*-problem anymore, because the second restriction becomes nonlinear due to the four restrictions following it. We will not give all the computations here, because they are lengthy and rather tedious. There is no interior solution of the first order conditions of an equivalent (transformed) problem. By checking all boundary (local) solutions, we found the global solution $q_8 = 1$; the connected reward to player B is 2.75. This boils down to player B always using his second action. So, player A possesses a strategy to hold his opponent's rewards to this amount, hence, $v^B \leq 2.75$.

Assume player B uses the second action at all stages of the play. In order to minimize his opponent's rewards, player A faces the following problem

$$\min_{q_2, q_4, q_6, q_8} 6q_2 + \frac{11}{2}q_4 + 3q_6 + \frac{11}{4}q_8$$

$$\text{s.t. } 1 = q_2 + q_4 + q_6 + q_8$$

$$0 = (1 - f_2)q_2 + (1 - f_4)q_4 - f_6q_6 - f_8q_8$$

$$f_3 = \left[\frac{1}{2} - \frac{11}{20}q_4 - \frac{11}{10}q_8\right]_+$$

$$f_4 = \left[\frac{1}{2} - \frac{11}{16}q_4 - \frac{11}{8}q_8\right]_+$$

$$f_7 = \left[\frac{1}{2} - \frac{11}{8}q_4 - \frac{11}{4}q_8\right]_+$$

$$f_8 = \left[\frac{1}{2} - \frac{11}{4}q_4 - \frac{11}{2}q_8\right]_+$$

$$0 \leq q_3, q_4, q_7, q_8.$$

Clearly, $q_8 = 1$ yields rewards equal to $\frac{11}{4}$; all other feasible rewards involve $q_8 < 1$ yielding a reward strictly higher than $\frac{11}{4}$. So, $v^B = 2.75$. ∎

6 Conclusions

We have expanded the framework of Small Fish Wars introduced in Joosten [2007a] and extended in Joosten [2007b,c, 2010a] by allowing stochasticity in the transition structure. We investigated three variants of randomness with increasing generality. The first one deals with transition probabilities which do not depend on the actions taken by the agents, but may depend on the state the system is in currently. The second variant treats transition
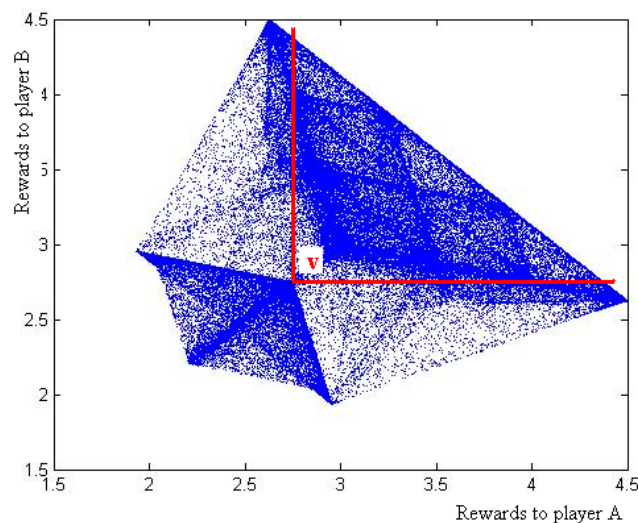
Figure 7: All jointly-convergent pure-strategy rewards to the 'north-east' of the two red lines can be supported by an equilibrium using threats. These equilibrium rewards can only be obtained if the agents use the action-pairs damaging the resource rarely.

probabilities depending on the actions taken by both agents and on the current state. These two variants are new for Small Fish Wars, yet fairly standard in the theory of stochastic games.

The third innovation however, is novel not only with respect to Small Fish Wars, but extends stochastic games into a new dimension. We allow the transition probabilities to depend on the actions taken by the agents, on the current state and on the history of the play until then. To the best of our knowledge, this aspect has not been examined before in the literature.

The inspiration for these innovations of the Small Fish Wars framework are mainly empirical. The first addition of randomness can be used to model the influence on the fish stock of weather, seasons, climate, water temperatures, or other random events beyond the control of the agents and without any relation to their catching behavior. The second innovation can additionally capture the influence of the players' behavior on the resource, especially differentiating between future effects on the fish stock of overfishing during periods of abundance and periods of scarcity. With this variant it is possible to induce a so-called social dilemma of the social trap type. The latter constitutes an empirical phenomenon which is far from exceptional in the exploitation of common pool resource systems.[7]

—————————

[7]The observation that social dilemmas occur frequently in real-life common pool resource systems by no means implies that the tragedy of the commons is ubiquitous. On

The third extension of Small Fish Wars with stochasticity allows to incorporate phenomena usually connected to low density of the resource. The model links past catching behavior to current transition probabilities. We allow that the state yielding bad outcomes becomes absorbing *temporarily*. If such a case arises, even if the agents turn to exploitation policies which are environmentally sound, it may take a non-trivial amount of time before the first transition to a state yielding higher outcomes occurs. In that manner we capture a poaching pit or hysteresis (cf., e.g., Bulte [2003]).

We solve these three extensions of the Small Fish War framework with methods introduced in Joosten, Brenner & Witt [2003] inspired by the literature on Folk Theorems for stochastic games e.g., Thuijsman & Vrieze [1998], Joosten [1996,2005], Schoenmakers, Flesch & Thuijsman [2002] and Schoenmakers [2004]. We obtain a continuum of rewards which can be supported by equilibria involving threats.

The good news is that in all variants, a 'tragedy of the commons' can be averted by sufficiently patient rational agents maximizing their utilities non-cooperatively. Almost all equilibrium rewards yield more than the amounts associated to the permanent ruthless exploitation of the resource. Pareto optimal equilibrium rewards always correspond to strategy pairs involving a considerable amount of restraint on the part of the agents and the associated rewards are notably better than no-restraint rewards.

The circumstance that in the example featuring endogenous transition probabilities which was analyzed rather completely for expository purposes, the threat point coincides with the no-restraint rewards is remarkable, but exceptional. In general this connection does not hold, in particular if the stage payoffs in *Low* are very small to the ones in *High*. It is also not generic that the state *Low* becomes absorbing temporarily.

We wanted to present a fairly tractable model and to economize on notations. Our first measure was to keep the fish stock fixed essentially yet stochastic, i.e., the variation in stock sizes and catches is only due to random effects captured by the transition probabilities. With respect to the original set-up of Small Fish Wars, i.e., with endogenous stage payoffs, we intend to combine the (possibly endogenous) stochastic effects examined and elaborated on here with changing fish stocks as in the original set-up.

Our second measure was to use three 'twos': two states, two players and two actions. Two distinct states allow to model the kind of transitions we had in mind. In order to capture additional phenomena observed in real-life systems, such as seasonalities or other types of correlations, a larger number of states may be required. Two agents are minimally required to model strategic interaction. We did not take more agents because complexities

---

the contrary, many societies locked into social dilemmas have found multiple designs to counteract the detrimental effects of the system and selfish myopic behavior (cf., e.g., Ostrom *et al.* [1994], Janssen & Ostrom [2001]).

in notation and representation arise, hardly justified by an added value. We used symmetry out of efficiency considerations. Another self-imposed limitation was the rather low number of stage-game actions. For applied work, more levels or dimensions of restraining measures may be necessary. For instance, when considering three types of nets, two types of boats, two options for the duration of fishing seasons combined with three levels of quota as measures for sustainable management of the resource, 36 actions would be a logical consequence. Adding states, (asymmetric) players or actions changes nothing to our approach conceptually.

The type of agent that we regard as being modeled, is not the individual fisherman, but rather countries, regions, villages or cooperatives of fishermen. It is debatable whether the latter types care for the future sufficiently to induce sustainability (see e.g., Ostrom [1990], Ostrom *et al.* [1994] for optimistic views), but individual fishermen's preferences seem too myopic (cf., e.g., Hillis & Wheelan [1994]). In Joosten [2007b], we briefly treat myopic time preferences as well, Joosten [2010b] discusses various social dilemmas arising within one and the same Small Fish War due to different levels of patience of the agents.

A modeling choice later on will be whether to model Allee effects with the set-up used in Joosten [2007b,c, 2010a] or with the present one. Adding a third state to the system which has only in-going transitions beyond a certain threshold and no out-going ones, together with zero stage-payoffs may do the trick of incorporating an Allee effect as well. Further research should also show whether time lags might assist in approximating the real-life phenomena aimed at here, more closely. Finally, splitting populations up into cohorts with different levels of fecundity, with the most fertile ones being hunted, may be another addition to the model worth exploring.

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.
Armstrong MJ, P Connolly, RDM Nash, MG Pawson, E Alesworth, PJ Coulahan, M Dickey-Collas, SP Milligan, M O'Neill, PR Witthames & L Woolner, 2001, An application of the annual egg production method to estimate spawning biomass of cod (Gadus morhua L.), plaice (Pleuronectes platessa L.) and sole (Solea solea L.) in the Irish Sea, *ICES J Mar Sci* **58**, 183-203.
Billingsley P, 1986, "Probability and Measure", John Wiley & Sons, New York, NY.
Blackwell D, 1962, Discrete dynamic programming, *Ann Math Stat* **33**, 719-726.

Bulte EH, 2003, Open access harvesting of wildlife: the poaching pit and conservation of endangered species, *Agricultural Econ* **28**, 27-37.

Courchamp F, E Angulo, P Rivalan, RJ Hall, L Signoret, & Y Meinard, 2006, Rarity value and species Extinction: The anthropogenic Allee effect. *PLoS Biol* **4**, 2405-2410.

Cross JG & MJ Guyer, 1980, "Social Traps", University of Michigan Press, Ann Arbor, Michigan.

Dawes RM, 1980, Social Dilemmas, *Annual Rev Psychology* **31**, 169-193.

Flesch J, 1998, "Stochastic Games with the Average Reward", PhD Thesis Maastricht University, ISBN 90-9012162-5.

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

Hall RJ, EJ Milner-Gulland & F Courchamp, 2008, Endangering the endangered: The effects of perceived rarity on species exploitation, *Conservation Letters* **1**, 75-81.

Hamburger H, 1973, N-person prisoner's dilemma, *J Math Psychology* **3**, 27-48.

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

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

Hillis JF, & BJ Wheelan, 1994, Fisherman's time discounting rates and other factors to be taken into account in planning rehabilitation of depleted fisheries, in: "Proc of the 6th Conf of the Intern Inst of Fisheries Econ and Trade" (M Antona, J Catanzano, JG Sutinen, eds.), IIFET-Secretariat Paris, *pp.* 657-670.

Hordijk A, OJ Vrieze & GL Wanrooij, 1983, Semi-Markov strategies in stochastic games, *Int J Game Theory* **12**, 81-89.

Hutchings JA, 2000, Collapse and recovery of marine fishes, *Nature* **406**, 882-885.

Janssen MA & E Ostrom, 2001, Critical factors that foster local self-governance of common pool resources: The role of heterogeneity, mimeo.

Joosten R, 1996, "Dynamics, Equilibria, and Values", Ph-D thesis no. **96-37**, Faculty of Economics & Business Administration, Maastricht University.

Joosten R, 2005, A note on repeated games with vanishing actions, *Int Game Theory Rev* **7**, 107-115.

Joosten R, 2007a, Small Fish Wars: A new class of dynamic fishery-management games, *ICFAI J Managerial Econ* **5**, 17-30.

Joosten R, 2007b, Small Fish Wars and an authority, in: "The Rules of the Game: Institutions, Law, and Economics" (A Prinz, AE Steenge & J Schmidt, eds.), *pp.* 131-162.

Joosten R, 2007c, Patience, Fish Wars, rarity value & Allee effects, *Papers Econ & Evol* **#0724**, Max Planck Institute of Economics, Jena.

Joosten R, 2010a, Strong and weak rarity value in Small Fish Wars, acceptance pending (improved and adapted version of Joosten [2007c]).

Joosten R, 2010b, Social dilemmas, time preferences and technology adoption in Small Fish Wars, mimeo.

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

Joosten R, H Peters & F Thuijsman, 1995, Unlearning by not doing: repeated games with vanishing actions, *Games Econ Behav* **9**, 1-7.

Kelly CJ, EA Codling & E Rogan, 2006, The Irish Sea cod recovery plan: some lessons learned, *ICES J Marine Science* **63,** 600-610.

Komorita SS & CD Parks, 1994, "Social Dilemmas", Brown & Benchmark's Social Psychology Series, Brown & Benchmark, Madison WI.

Kurlansky M, 1998, "Cod: A Biography of the Fish That Changed the World", Vintage, Canada.

Levhari D & LJ Mirman, 1980, The great fish war: An example using a dynamic Cournot-Nash solution, *Bell J Econ* **11**, 322-334.

Liebrand WBG, 1983, A classification of social dilemma games, *Simulation & Games* **14**, 123-138.

Olson M, 1965, "The logic of collective action", Harvard Press, Cambridge, MA.

Oosthuizen E & N Daan, 1974, Egg fecundity and maturity of North Sea cod, gadus morhua, *Netherlands J Sea Res* **8**, 378-397.

Ostrom E, 1990, "Governing the Commons", Cambridge University Press, Cambridge, UK.

Ostrom E, R Gardner & J Walker, 1994, "Rules, Games, and Common-Pool Resources", Michigan University Press, Ann Arbor, MI.

Platt J, 1973, Social traps, *American Psychologists* **28**, 641-651.

Rose GA, IR Bradbury, B deYoung, SB Fudge, GL Lawson, LGS Mello, D Robichaud, G Sherwood, PVR Snelgrove & MJS Windle, 2008, Rebuilding Atlantic Cod: Lessons from a Spawning Ground in Coastal Newfoundland in: "Resiliency of gadid stocks to fishing and climate change" (GH Kruse, K Drinkwater, JN Ianelli, JS Link, DL Stram, V Wespestad, D Woodby, eds.), 24th Lowell Wakefield Fisheries Symposium, ISBN 978-1-56612-126-2, *pp.* 197-219

Russ GR & AC Alcala, 1996, Marine reserves: rates and patterns of recovery and decline of large predatory fish, *Ecol Applic* **6**, 947-961.

Russ GR & AC Alcala, 2004, Marine reserves: long term protection is required for full recovery of predatory fish populations, *Conservation Ecol* **138**, 622-627.

Schelling TC, 1978, "Micromotives and Macrobehavior", WW Norton, NY.

Schoenmakers GM, 2004, "The Profit of Skills in Repeated and Stochastic Games", Ph-D Thesis Maastricht University, ISBN 90 90184473.

Schoenmakers GM, 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.

Shertzer KW & MH Prager, 2006, Delay in fishery management: diminished yield, longer rebuilding, and increased probability of stock collapse, *ICES J Marine Science* **64**, 149-159.

Skyrms B, 2004, "The Stag Hunt and the Evolution of Social Structure", Cambridge University Press, Cambridge, UK.

Tegner MJ, 1992, Brood-stock transplants as an approach to abalone stock enhancement, in: "Abalone of the World: Biology, Fisheries, and Culture" (SA Shepherd, MJ Tegner and SA Guzman del Proo, eds.), Blackwell, Oxford, UK, *pp.* 461-473.

Thuijsman F, 1992, "Optimality and Equilibria in Stochastic Games", *CWI-Tract* **82**, Centre for Mathematics and Computer Science, Amsterdam.

Thuijsman F & OJ Vrieze, 1998, The power of threats in stochastic games, in: "Stochastic and Differential Games, Theory and Numerical Solutions" (M Bardi, TES Raghavan & T Parthasarathy, eds.), Birkhauser, Boston, *pp.* 343-358.

Van Damme EEC, 1992, "Stability and Perfection of Nash Equilibria, Springer, Berlin.

Vieille N, 2000a, Two-player stochastic games I: A reduction, *Israel J Math* **119**, 55-91.

Vielle 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.