Semi-Stationary Equilibria in Dynamic Games [⋆]

Mitri Kitti

Department of Economics, Aalto University School of Economics, P.O. Box 21240, FI-00076 Aalto, Finland

---

Abstract

This paper introduces stationary simple strategies and related stationary penal codes in the spirit of Abreu (1986, 1988) for dynamic games with perfect monitoring. The class of equilibria that these strategies support is defined—they are called semi-stationary equilibria. The main results are necessary and sufficient conditions for strategies, or the corresponding penal codes, which give the minimal payoffs to players among semi-stationary equilibria. Results are applied to resource extraction games.

Key words: dynamic game, subgame-perfect equilibrium, penal code, simple strategy, stationary strategy, resource extraction

JEL Classification: C61, C72, C73, Q20

---

1 Introduction

Dynamic games provide a framework for modeling behavior in situations where agents' actions influence the environment in which they interact, and this influence is captured by a state variable. This paper deals with both deterministic and stochastic discounted dynamic games with discrete time, infinite time horizon, and stationary payoffs and dynamics. The paper follows the line of research dating back to Shapley (1953) that aims to generalize the ideas of dynamic programming to games.

[⋆] Funding from the Academy of Finland is gratefully acknowledged.
Email address: `mitri.kitti@hse.fi` (Mitri Kitti).

The main questions related to dynamic games, whether they are stochastic or deterministic, are the nature of solutions and their existence. In this paper the emphasis is on addressing the first issue by introducing a new class of strategies resulting to stationary play after any history. These strategies are called semi-stationary.

The most important solution concept for dynamic games is subgame-perfect equilibrium. However, a significant part of literature deals with Markov equilibria which are subgame-perfect equilibria defined by strategies that are conditional on the state variable instead of the whole history of play. Stationary Markov strategies belong to the class of semi-stationary strategies. This means that existence results for stationary Markov equilibria in discounted dynamic games are also valid for semi-stationary strategies, see, e.g., Duffie et al. (1994), Horst (2005), Nowak (2007), Nowak and Raghavan (1992), and Parthasarathy and Sinha (1989).

When a game has Markov equilibria, there are usually many of them. Moreover, knowing one equilibrium does not tell much about the other. In particular, to find the players' smallest or largest Markov equilibrium payoffs can be very difficult. On the contrary, semi-stationary strategies have the property that the ones giving the smallest payoffs to players, called extremal strategies, can be found without knowing all equilibria. Furthermore, once the smallest payoffs are found, we can use them in finding the largest payoffs.

In the context of repeated games, Abreu (1988) showed that all the subgame-perfect equilibria can be characterized using extremal equilibrium strategies, i.e., those strategies that yield the smallest equilibrium payoffs to players. To be more specific, if a path of actions is induced by a subgame-perfect equilibrium strategy, then a simple strategy in which deviations from this path are punished with reverting to extremal strategy of a deviator is subgame perfect and vice versa. Polasky et al. (2006) and Mason et al. (2009) have recently analyzed this type of strategies in dynamic games arising in environmental models.[1] In this framework a subgame perfect simple strategy can be interpreted as a self-enforcing agreement.

This paper introduces stationary simple strategies for dynamic games with perfect monitoring in the spirit of Abreu (1986, 1988). Stationary simple strategies are defined by stationary policies that players follow initially and penal codes corresponding to deviations from the ongoing policies. A stationary penal code determines a set of stationary policies that are followed after deviations. The idea of a penal code is close to concepts such as recursively supported equilibria (Cave, 1987) and trigger-strategy equilibria (Benhabib and Radner, 1992) that have been introduced for models of exploiting a common resource pool. Sim-

---

[1] Abreu (1986, 1988), Polasky et al. (2006), and Mason et al. (2009) do not make assumptions on stationarity.

ple stationary strategies, recursively supported equilibria, and trigger-strategy equilibria all belong to semi-stationary equilibria, but there are also other semi-stationary equilibria. However, all of them are supported by simple stationary strategies.

The restriction to stationary strategies, and more generally to semi-stationary strategies, can be motivated by two reasons. First, they are not overly complicated as they prescribe stationary behavior. Semi-stationary strategies are not the most simple ones but they have the benefit that we can characterize equilibria using extremal strategies—a property that stationary Markov strategies lack. The second argument, adopted from Duffie et al. (1994), is that equilibria prescribing non-stationary behavior require sophisticated coordination that may be implausible.

Semi-stationary equilibria have the property that each of them is supported by an extremal stationary penal code. These are the equilibrium penal codes that define stationary simple strategies which give the minimal equilibrium payoffs to players. It is shown that these penal codes result to payoff functions that satisfy a rather simple fixed-point equation which makes them computationally tractable. This paper restricts to pure strategies but generalizing the main concepts and results to mixed strategies is straightforward. Resource extraction games are studied as an application.

The paper is structured as follows. Stationary simple strategies are defined in Section 2. Section 3 introduces the concept of extremal stationary penal code. Semi-stationary strategies and their relation to extremal stationary penal codes is analyzed in Section 4. Application to resource extraction games is presented in Section 5. Finally, the results are discussed in Section 6.

2 The Model

2.1 Dynamic Games

There are n players indexed with $i \in I = \{1, \ldots, n\}$. The players' available actions at state $x \in X$ are $Y_i(x)$, $i \in I$. The correspondence of feasible action profiles is denoted as

$$Y(x) = Y_1(x) \times \cdots \times Y_n(x), \text{ and } Y = \cup_{x \in X} Y(x).$$

The state evolves according to a dynamical system

$$x^{k+1} = f\left(y^k, x^k, w^k\right), \; k = 0, 1, \ldots,$$

where f is a function from $Y \times X \times W$ to the set of states X.[2] Here W is the set of random disturbances affecting the evolution of the state. To be more specific, after players have simultaneously chosen their actions y^k, the disturbance w^k is realized and the state transition takes place. The random disturbances w^k, $k \geq 0$, are identically distributed. The probabilities are given by $p(\mathrm{d}w^k | y^k, x^k)$ defined on a probability space $(W, \mathcal{F})$.

In each stage the players receive payoffs determined by their utility functions $u_i : Y \times X \times W \mapsto \mathbb{R}$, $i \in I$. Perfect monitoring is assumed; the players observe perfectly each others' actions as well as the disturbances. The duration of the game is infinite. A history in stage k is denoted as h^k and it is defined recursively as follows: $h^0 = x^0$, $x^0 \in X$, and for $k \geq 1$ the history is $h^k = h^{k-1} \cup \{y^{k-1}, w^{k-1}\}$. Note that the state after history h^k is fully determined from the dynamical system governing the state transition. Therefore, other states than the initial one are not included in the history.

The set of all possible histories in stage k for any initial state is denoted as H^k. Moreover, $x(h^k)$ is the state that has been reached after history $h^k \in H^k$. A strategy for player i is a sequence of functions $(\sigma_i^0, \sigma_i^1, \ldots)$ where σ_i^k maps any $h^k \in H^k$ to $Y_i(x(h^k))$. A strategy profile is denoted as $\sigma = (\sigma_i, \ldots, \sigma_n)$ and σ_{-i} is the collection of strategies of other players than player i.

The players maximize their expected discounted utilities. The discount factor of player i is $\delta_i \in (0, 1)$. Unlike in most papers on discounted dynamic games, the discount factors need not be the same. The objective of player i' is to maximize

$$\lim_{m\to\infty} \mathbb{E}\left[\sum_{k=0}^{m} \delta_i^k u_i\left(\sigma_i^k(h^k), \sigma_{-i}^k(h^k), x(h^k), w^k\right)\right]$$

over strategies $\sigma_i = (\sigma_i^0, \sigma_i^1, \ldots)$. The expectation is over disturbance sequences $w^0, w^1, \ldots$. Moreover, the limit in the above formula is assumed to exist. The conditions under which this is the case for the class of strategies considered in this paper will be discussed later.

When the players follow σ, the payoff for player i beginning from k-period history h^k is

$$U_i\left(\sigma, h^k\right) = \lim_{m\to\infty} \mathbb{E}\left[\sum_{j=0}^{m} \delta_i^j u_i\left(\sigma^{k+j}(h^{k+j}), x(h^{k+j}), w^{j+k}\right)\right], \tag{1}$$

Subgame perfection is now defined in the usual manner.

Definition 1. Strategy profile σ is a subgame-perfect equilibrium (SPE) if

---

[2] Even when there is no physical state, it is possible to define state variables, e.g., in a repeated game we can set $f(y) = y$, i.e., $X = Y$. The results of this paper are valid also for this kind of choices of state variables.

for all $k \geq 0$ and $i \in I$ we have

$$U_i\left(\sigma, h^k\right) \geq U_i\left(\sigma_i', \sigma_{-i}, h^k\right)$$

for all strategies σ_i' and for all $h^k \in H^k$.

If we have $U_i(\sigma, h^k) \geq U_i(\sigma_i', \sigma_{-i}, h^k) - \varepsilon$ for $\varepsilon > 0$ in the above definition, we say that σ is an ε-subgame-perfect equilibrium.

2.2 Stationary Simple Strategies and Penal Codes

As in the literature on dynamic programming, if the player's actions at each stage are determined by a function μ_i that maps $x \in X$ into $Y_i(x)$, the player follows a stationary policy defined by μ_i. In the following μ_i is called a stationary policy for player i. Stationary policy profiles are composed of stationary policies, i.e., $\mu : X \mapsto Y$ is a stationary policy profile when $\mu(x) \in Y(x)$ for all $x \in X$. Stationary policy profiles are assumed to belong to a space $M = M_1 \times \cdots \times M_n$, i.e., $\mu_i \in M_i$, $i \in I$. By $\mu_{-i} \in M_{-i}$ we denote the policies of other players than player i. The assumptions that we need for M will be commented in the end of this section.

Simple strategies are composed of policy profiles which define the players' actions in all contingencies; what is played initially, what is played after one of the players unilaterally deviates from the initial play, and what is played after deviations from these policy profiles. Simultaneous deviations are ignored. The policy profiles that are played after deviations form a stationary penal code.

Definition 2. A strategy profile $\sigma(\mu^0, \mu^1, \ldots, \mu^n)$ is a stationary simple strategy profile if players start by following $\mu^0 \in M$, and if player i was the last to deviate unilaterally from the ongoing policy, the players follow $\mu^i \in M$. The collection $\{\mu^1, \ldots, \mu^n\}$ is called a stationary penal code.

In the following Σ^s is the set of stationary simple strategies that are subgame perfect. Moreover, $V(\Sigma^s)$ denotes the corresponding set of payoff functions, i.e., $v \in V(\Sigma^s)$ means that there is $\sigma \in \Sigma^s$ such that player i's continuation function is $v_i(x)$ when the players follow σ. A stationary penal code $\{\mu^1, \ldots, \mu^n\}$ is said to be subgame-perfect if $\sigma(\mu^i, \mu^1, \ldots, \mu^n) \in \Sigma^s$ for all $i \in I$. The set of all stationary penal codes that result to subgame-perfect equilibria is denoted as Σ^p. Notice that, penal codes themselves are not strategies but they can be used in forming simple strategies. In particular $\sigma(\mu^i, \mu^1, \ldots, \mu^n)$, $i \in I$, are simple strategies defined by the stationary penal code $\{\mu^1, \ldots, \mu^n\}$. The following result is evident because stationary Markov strategies define stationary simple strategies.

Remark 1. $\mu \in M$ is a policy profile corresponding to a stationary Markov equilibrium if and only if $\sigma(\mu, \dots, \mu)$ is SPE.

It follows that if a game has a stationary Markov equilibrium, then it has a stationary simple equilibrium, too. The existence of Markov strategies is, however, a non-trivial question, particularly for pure strategies. For recent papers on the existence of pure strategy Markov equilibria see, e.g., Horst (2005) and Nowak (2007).

Mixed strategies are not considered in this paper. The main concepts, however, generalize to these strategies when instead of defining policies as mappings with values in $Y_i(x)$, we define them as mappings into distributions over $Y_i(x)$ and take the expectation in (1) over sequences of disturbances and actions. Recall that that players' actions are assumed to be observable. For mixed strategies we would need the assumptions that players observe the distributions used in random selections of actions. However, this is not necessarily realistic in many cases. Note also that when actions are observable, restricting to Markov strategies can be questioned. On the other hand, when only the state variable is observed, Markov strategies are a natural choice.

The notion of simple strategy is adopted from Abreu (1986, 1988) who introduced it for repeated games without the assumption on stationarity, which plays a crucial role in this paper. In a repeated game a stationary simple strategy is a trigger-type strategy in which a unilateral deviation from a predetermined action profile y^0 leads to another action profile y^i when the player i was the deviator. In the following, we give examples of stationary simple strategies in a deterministic dynamic game with two states. We also make some comparisons to Markov strategies to clarify matters.

Example 1. Let us assume that there are two players and two states x_1 and x_2, $Y_1(x_1) = Y_1(x_2) = \{a_1, a_2\}$, $Y_2(x_1) = Y_2(x_2) = \{b_1, b_2\}$, and there is a deterministic state transition function f such that $f(a_2, b_2, x_1) = x_2$, $f(y, x_1) = x_1$ for $y \neq (a_2, b_2)$, $f(a_2, b_2, x_2) = x_2$, $f(y, x_2) = x_1$ for $y \neq (a_2, b_2)$. The payoffs are as below where $*$ signifies a state transition.

x_1	b_1	b_2
a_1	$(4, 4)$	$(0, c)$
a_2	$(c, 0)$	$(1, 1)^*$

x_2	b_1	b_2
a_1	$(0, 0)^*$	$(-4, c-4)^*$
a_2	$(c-4, -4)^*$	$(-3, -3)$

The number c is assumed to be greater than one. For $c > 4$ the stage games are variations of prisoners' dilemma. The twist is that if in state x_1 players defect, i.e., choose (a_2, b_2), they end up to state x_2 in which the possible gains are smaller for both of them. On the other hand, they get back to state x_1 if at least one of the players cooperates in state x_2.

An example of a stationary simple strategy is determined by the stationary policies

$$\mu^0(x_i) = (a_1, b_1),\ \mu^1(x_i) = (a_1, b_2),\ \mu^2(x_i) = (a_2, b_1),\ i = 1, 2.$$

The players start by cooperation in both states, and if player i deviates they switch to play according to policy profile μ^i. In this example $\{\mu^1, \mu^2\}$ is a stationary penal code. The policies μ^i_j, $j, i \in I$, can be regarded as Markov strategies for the players. Whether μ^i, $i \in I$, are Markov equilibria depend on c and the discount factors. The policy profile μ^0 is a Markov equilibrium for $c \leq 4$ and all $\delta_i \in (0, 1)$, $i = 1, 2$. Policies μ^1 and μ^2 cannot be Markov equilibria for any $\delta_i \in (0, 1)$, $i = 1, 2$, when $c < 4$ because in that case deviating to (a_1, b_1) is profitable in both states. It can also be observed that there are not necessarily any Markov equilibria in pure strategies for $c > 4$ depending on discount factors. However, $\mu^0(x_1) = (a_2, b_2)$, $\mu^0(x_2) = (a_2, b_1)$ (or $\mu^0(x_2) = (a_1, b_2)$ alternatively), becomes an equilibrium when $\delta_1 \leq 1/4$ and $\delta_2 \geq 1/4$.

Before going into the analysis of stationary penal codes let us briefly discuss the conditions under which the expression in (1) is well-defined for stationary simple strategies. The following structural assumption should hold for the class of policy profiles M.

(A1) The below limit exists for all $i \in I$, $\mu \in M$, and $x^0 \in X$:

$$\lim_{m \to \infty} \mathbb{E} \left[\sum_{k=0}^{m} \delta_i^k u_i \left(\mu(x^k), x^k, w^k \right) \right], \tag{2}$$

where x^k satisfy $x^{k+1} = f(\mu(x^k), x^k, w^k)$ for all $k = 0, 1, \ldots$.

As explained in detail by Bertsekas and Shreve (1996), there are basically two approaches to guarantee (A1). The first is to extend the notion of integral and assume that the expectation is defined as an outer integral. The second is to make measurability assumptions which guarantee that (2) is well-defined. In the latter case suitable assumptions are provided by the Borel-semianalytic framework:

(1) X, Y, and W are Borel spaces,
(2) $p(\mathrm{d}w|y, x)$ is a Borel measurable stochastic kernel on W for all $y \in Y$ and $x \in X$,
(3) $\Gamma = \{(y, x, w) :\ x \in X,\ y \in Y(x),\ w \in W\}$ is analytic in $Y \times X \times W$,
(4) f is a Borel measurable function from $Y \times X \times W$ to X,
(5) u_i, $i \in I$, are upper semianalytic functions from Γ to $\mathbb{R}$,
(6) M consists of universally measurable functions for which $\mu(x) \in Y(x)$ for all $x \in X$.

In the following sections we shall simply assume that (A1) holds.

2.3 Characterization of Stationary Equilibrium Penal Codes

Before going into the analysis some notation is needed. The $n \times n$ continuation functions corresponding to $\mu^1, \dots, \mu^n$ are denoted as $v(\mu^1, \dots, \mu^n)$. Each component $v_j^i(\mu^1, \dots, \mu^n)$, $j, i \in I$, gives the payoff to player j when μ^i is played. When omitting the subscript, i.e., denoting v^i, we refer to function from X to $\mathbb{R}^n$ that has component functions corresponding to the payoffs of all players when they follow the stationary policy profile μ^i. When we refer to payoffs given by the policy profile μ^i we also denote $v(\mu^i)$, i.e., $v^i(\mu^1, \dots, \mu^n) = v(\mu^i)$. If $v_j(\mu^i)(x) \leq v_j(\nu^i)(x)$ for all $x \in X$ and $j \in I$, we denote $v(\mu^i) \leq v(\nu^i)$. Because we may have genuinely interesting policies that lead to arbitrarily small or large payoffs at some states, the functions v_j^i, $i, j \in I$, are assumed to be extended real-valued functions, i.e., they may take values ∞ or $-\infty$.

To shorten the notation let us denote

$$T_j\left(y, x, v_j^i\right) = \mathbb{E}\left[u_j\left(y, x, w\right) + \delta_j v_j^i(f(y, x, w))\right],$$

where v_j^i is a continuation value function for player j, i.e., $v_j^i : X \mapsto \mathbb{R}$. The expectation is over w. Observe that for $v = v(\mu^1, \dots, \mu^n)$ we have $v_j^i(x) = T_j(\mu^i(x), x, v_j^i)$ for all $j, i \in I$. Continuation value functions are assumed to belong to a suitable function space F. For example, when the model is Borel-semianalytic, F is the space of functions which have upper semianalytic components. The following property is assumed.

(A2) The expected value in the definition of $T_j(\mu^i(x), x, v_j^i)$ exists for all $i, j \in I$, $\mu \in M$, $x \in X$, and $v \in F$.

As in dynamic programming, even though taking expectations and limits were well-defined, more assumptions are needed in order to obtain necessary and sufficient conditions for an equilibrium. In particular we need the one-shot deviation principle. Without making any particular assumptions for utility functions, it will be assumed that the one-shot deviation principle holds.[3] In other words, given that other players punish deviations from μ^i according to the penal code $\{\mu^1, \dots, \mu^n\}$, it is optimal for player j to follow μ^i if and only if following μ^i gives a higher payoff than deviating from it in one stage and then following μ^j as prescribed by the penal code strategy.

---

[3] One-shot deviation principle holds for Borel-semianalytic models when payoffs are bounded below, see, e.g., Proposition 9.13 in Bertsekas and Shreve (1996). When the payoffs are not bounded below it may happen that a suboptimal policy satisfies the one-shot deviation property.

Formally, one-shot deviation principle is defined by assuming that player j is choosing the optimal strategy given that the other players follow $\mu^i_{-j} \in M_{-j}$, if player j has played according to $\mu^i_j \in M_j$, and otherwise they follow $\mu^j_{-j} \in M_{-j}$. The one-shot deviation principle gives a necessary and sufficient condition for the optimality of the strategy in which player j chooses to play according to μ^i_j and follow μ^j_j after he has deviated. In that case the pair μ^i_j, μ^j_j is optimal for player j.

(A3) For all $j \in I$ and $\mu^i_{-j}, \mu^j_{-j} \in M_{-j}$ the pair μ^i_j, $\mu^j_j \in M_j$ is optimal for player j if and only if

$$T_j\left(\mu^k(x), x, v_j(\mu^k)\right) \geq \sup_{y_j \in Y_j(x)} T_j\left(y_j, \mu^k_{-j}(x), x, v_j(\mu^j)\right) \; \forall x \in X, \; k = i, j.$$

Note that for $k = j$ the above inequality holds with equality, i.e., μ^j_j is the optimal policy of player j when the other players follow μ^j_{-j}.

Corresponding to a continuation payoff v and player i, the set of incentive compatible action profiles at state $x \in X$ consists of $y \in Y(x)$ which satisfy

$$T_j\left(y, x, v^i_j\right) \geq \sup_{y_j \in Y_j(x)} T_j\left(y_j, y_{-j}, x, v^j_j\right) \text{ for all } j \in I. \tag{3}$$

This set is denoted as $IC_i(v)(x)$. If $y \in IC_i(v)(x)$, then all players prefer playing y and receiving continuation payoffs v^i_j, $j \in I$, to deviating and then receiving v^j_j, $j \in I$. The following result is a direct consequence of (A3).

Lemma 1. *Under (A1)–(A3),* $\{\mu^1, \ldots, \mu^n\} \in \Sigma^p$ *if and only if* $\mu^i \in M$ *and* $\mu^i(x) \in IC_i(v)(x)$ *for* $v = v(\mu^1, \ldots, \mu^n)$ *and for all* $i \in I$ *and* $x \in X$.

Recall that v^i_j is assumed to be extended real-valued functions. The value $-\infty$ would mean that a players utility function u_i is not bounded below in which case (A3) does not necessarily hold. In some cases, however, (A3) may hold even when u_i, $i \in I$, were unbounded from below.

When dealing with pure strategies it may happen that there are no stationary equilibrium penal codes. Indeed, it is rather easy to form a repeated game where there are no SPE in pure strategies. The following proposition, however, gives necessary and sufficient conditions for the existence of stationary equilibrium penal codes.

Proposition 1. *Under (A1)–(A3), there is a stationary equilibrium penal code if and only if there are* $v \in F$ *and* $\mu^i \in M$, $i \in I$, *such that* $\mu^i(x) \in IC_i(v)(x)$ *for all* $i \in I$, $x \in X$, *and*

$$T_i\left(\mu^i(x), x, v^i_i\right) \leq v^i_i(x), \; T_j\left(\mu^i(x), x, v^i_j\right) \geq v^i_j(x) \; \forall i, j \in I, \; j \neq i, \; x \in X. \tag{4}$$

Proof. If $\{\mu^1, \ldots, \mu^n\}$ is a stationary penal code, and $v_j^i = v_j(\mu^i)$, then $\mu^i(x) \in IC_i(v)(x)$ by Lemma 1, and the inequalities in (4) hold with equality. Hence, if there is a stationary equilibrium penal code, the conditions of Proposition 1 hold.

For the other direction of the proof we need a result that the following inequalities

$$\bar{v}_i^i \leq v_i^i \text{ and } \bar{v}_j^i \geq v_j^i \text{ for all } i \in I, \ j \neq i, \tag{5}$$

and $\mu^i(x) \in IC_i(v)(x)$ imply that $\mu^i(x) \in IC_i(\bar{v})(x)$ for all $i \in I$ and $x \in X$. Let us now show this. The second inequality in (5) gives $T_j(\mu^i(x), x, \bar{v}_j^i) \geq T_j(\mu^i(x), x, v_j^i)$. From the first one we obtain

$$\sup_{y_j \in Y_j(x)} T_j\left(y_j, y_{-j}, x, \bar{v}_j^j\right) \leq \sup_{y_j \in Y_j(x)} T_j\left(y_j, y_{-j}, x, v_j^j\right).$$

These two inequalities and the fact that $\mu^i(x) \in IC_i(v)(x)$ give us incentive compatibility $\mu^i(x) \in IC_i(\bar{v})(x)$;

$$\sup_{y_j \in Y_j(x)} T_j\left(y_j, \mu_{-j}^i, x, \bar{v}_j^j\right) \leq T_j\left(\mu^i(x), x, \bar{v}_j^i\right).$$

Next, we argue that

$$v_i(\mu^i) \leq v_i^i \text{ and } v_j(\mu^i) \geq v_j^i \text{ for all } i \in I, \ j \neq i. \tag{6}$$

Observe that $v_j(\mu^i)(x)$ is obtained as the limit of the iteration

$$v_{ji}^{k+1}(x) = \mathbb{E}\left[u_j\left(\mu^i(x), x, w\right) + \delta_j v_{ji}^k\left(f(\mu^i(x), x, w)\right)\right], \ k = 0, 1, \ldots$$

for $v_j^0 = v_j^i$. By induction argument, $v_{ii}^k \leq v_i^i$ and $v_{ji}^k \geq v_j^i$. It follows that the pointwise limit satisfies these inequalities as well, i.e., (6). Hence, we obtain $\mu^i(x) \in IC_i(\bar{v})(x)$ for $\bar{v}$ with $\bar{v}^i = v(\mu^i)$ for all $i \in I$. In other words, μ^i are incentive compatible when the continuation payoffs are given by $v(\mu^i)$. Hence, by Lemma 1 policy profiles $\mu^1, \ldots, \mu^n$ give us a stationary equilibrium penal code. □

Proposition 1 gives us means of verifying whether there are stationary simple equilibria. This is because $\Sigma^s \neq \emptyset$ is equivalent to $\Sigma^p \neq \emptyset$, i.e., there are stationary equilibrium penal codes if and only if there are stationary simple equilibria.

3 Extremal Stationary Penal Codes

This section introduces stationary penal codes that yield the smallest payoffs to players. These penal codes are called extremal stationary penal codes. As will be shown in Section 4 they can be used in supporting any strategy that leads to a play that is described by a stationary policy profile after all histories. The main results of this section are necessary and sufficient conditions for extremal stationary penal codes.

An extremal stationary simple strategy is an equilibrium strategy in which one of the players' payoff is minimized. An extremal stationary penal code is one that leads to extremal simple strategies.

Definition 3. Penal code $\{\mu^1, \dots, \mu^n\}$ is an extremal stationary penal code (ESPC) if for all $i \in I$ the strategy $\sigma(\mu^i, \mu^1, \dots, \mu^n)$ is extremal for player i, i.e., $v_i^i(\mu^0) \le v_i^i(\nu^0)$ for all $\sigma(\nu^0, \nu^1, \dots, \nu^n) \in \Sigma^s$ and $i \in I$.

The set of extremal stationary penal codes is denoted as $\mathrm{ext}(\Sigma^p)$.

The following lemma tells that an ESPC satisfies a necessary condition that is similar to the necessary and sufficient optimality condition in dynamic programming.

Lemma 2. *Let us assume that (A1)–(A3) hold. If* $\{\mu^1, \dots, \mu^n\} \in ext(\Sigma^p)$ *then* $v = v(\mu^1, \dots, \mu^n)$ *satisfies*

$$T_i\left(\mu^i(x), x, v_i^i\right) = \inf\left\{T_i\left(y, x, v_i^i\right) : y \in IC_i(v)(x)\right\}\ \forall i \in I. \tag{7}$$

Proof. If $\{\mu^1, \dots, \mu^n\} \in \mathrm{ext}(\Sigma^p)$, then $\mu^i(x) \in IC_i(v)(x)$ for all $i \in I$ by Lemma 1, and there is no other $y \in IC_i(v)(x)$ that gives a smaller value for $T(y, x, v_i)$. Otherwise we would have a contradiction with the extremality of the strategy. Hence, (7) holds. $\square$

The set of ε-incentive compatible actions can be defined by subtracting ε from the right hand side of (3) for all $i \in I$. Respectively, we need subtract ε from the right hand side of (7). In that case we obtain the above result for ε-subgame-perfect ESPCs. The rest of the paper deals with SPE strategies, although in most cases we could obtain the same results for ε-subgame-perfect strategies, too.

The result of Lemma 2 resembles the fact that in dynamic programming payoffs corresponding to an optimal policy satisfy Bellman equation. Unlike in dynamic programming, it is not obvious how to formulate the necessary condition (7) as a fixed point of a suitable operator. The reason is that if we fix

v and take the infimum in the right hand side of (7) for all $i \in I$, the other components, $j \neq i$, are not necessarily uniquely defined even when the infima were attained.

The condition (7) as such is only a necessary condition for an ESPC. The following result tells that an ESPC defines payoff functions v such that there are no other stationary penal codes that satisfy (7) and give smaller payoffs at all states than v_i^i. For this result we need assume that all extremal payoffs given by stationary simple strategies are attained. In the following we denote the closure of $V(\Sigma^s)$ in the topology of pointwise convergence as $\bar{V}(\Sigma^s)$. In other words, if $v \in \bar{V}(\Sigma^s)$, then there is a sequence of stationary simple strategies giving payoffs v^k such that $v^k(x)$ converge to $v(x)$ for all $x \in X$.[4]

(A4) If for some $i \in I$, $\bar{v} \in \bar{V}(\Sigma^s)$ satisfies $\bar{v}_i \leq v_i$ for all $v \in V(\Sigma^s)$, then there is $\sigma \in \Sigma^s$ that gives $\bar{v}$ when players follow σ.

An important observation on condition (A4) is that it assures that $v \in V(\Sigma^s)$ if and only if there is a policy profile that leads to equilibrium payoffs when it is supported with an ESPC. Namely, if $v \in V(\Sigma^s)$, then there is $\{\bar{\mu}^1, \ldots, \bar{\mu}^n\} \in \Sigma^p$ such that $\sigma(\mu, \bar{\mu}^1, \ldots, \bar{\mu}^n) \in \Sigma^s$. Due to (A4), there is $\{\mu^1, \ldots, \mu^n\} \in \Sigma^p$ such that $v_i(\bar{\mu}^i)(x) \geq v_i(\mu^i)(x)$ for all $i \in I$ and $x \in X$. It follows from (A3) that $\sigma(\mu, \mu^1, \ldots, \mu^n) \in \Sigma^s$.

Remark 2. Under (A1)–(A4), $v \in V(\Sigma^s)$ if and only if there is $\mu \in M$ and $\{\mu^1, \ldots, \mu^n\} \in \Sigma^p$ such that $v = v(\mu)$ and $\sigma(\mu, \mu^1, \ldots, \mu^n) \in \Sigma^s$.

Sufficient conditions for an ESPC under (A1)–(A4) are stated in Proposition 2. Recall that Proposition 1 gives a condition for the existence of stationary equilibrium penal codes. However, even if there are equilibria, there are not necessarily any ESPCs. This happens when the extremal payoffs are not reached with any strategy. Section 5 presents an example where this is the case. In this paper the focus is not on any specific conditions that would guarantee (A4). Observe, however, that when considering ε-subgame perfect equilibria we do not need (A4).

Proposition 2. *Under (A1)–(A4), $\{\mu^1, \ldots, \mu^n\} \in ext(\Sigma^p)$, if and only if $v = v(\mu^1, \ldots, \mu^n)$ satisfies*

$$\begin{aligned} &v_i^i(\{\nu^1, \ldots, \nu^n\}) \geq v_i^i, \text{ for all } i \in I \text{ and } \{\nu^1, \ldots, \nu^n\} \in \Sigma^p \\ &\text{that satisfy (7) for } v(\{\nu^1, \ldots, \nu^n\}). \end{aligned} \tag{8}$$

Proof. The necessity of (7) follows from Lemma 2. Observe also that if there is an other strategy that satisfies (7) and defines a continuation payoff $\tilde{v}$ such

---

[4] When X is a Borel space the pointwise convergence is needed for almost all $x \in X$.

that $\tilde{v}_i^i(x) \leq v_i^i(x)$ for all $i \in I$ and $x \in X$ with strict inequality for some $i \in I$ and $x \in X$, then we would have a contradiction with the extremality of the strategy.

Let us assume that $\{\mu^1, \dots, \mu^n\}$ satisfies (7) for $v = v(\mu^1, \dots, \mu^n)$. Clearly, $\sigma(\mu^i, \mu^1, \dots, \mu^n)$, $i \in I$, are subgame perfect. If there was another set of policy profiles defining subgame-perfect strategies and giving a smaller value for some $v_i^i(x)$ for some $x \in X$, then we would have a contradiction with (8). □

In practice, the main challenge in testing whether a stationary penal code which satisfies (7) is an ESPC comes from the condition (8). One resolution is provided by the following result, where the security level refers to the minmax payoff function of a player, i.e., the payoff that a player can guarantee himself or herself when the other players are minimizing this outcome. The result follows directly from (8). Observe that when a stationary penal code leads to a player's security level then this penal code defines a strategy that leads to the smallest subgame perfect payoff that the player can get. As seen in the following section, generally extremal penal codes lead to least payoffs among semi-stationary strategies.

Corollary 1. *When a stationary penal code satisfies (7) and μ^i gives the security level of player i, then the penal code is extremal for player i.*

In the examples of this paper ESPCs are determined by strategies that give players their security levels. In general this need not be the case. Counter examples can be easily constructed, e.g., for repeated games by choosing payoffs such that security strategies are not Nash equilibria of the stage game.

Example 2. Let us continue the first example. The aim is to show that policies μ^1 and μ^2 introduced previously form an ESPC for certain range of discount factors. First, observe that the continuation payoffs corresponding to $\{\mu^1, \mu^2\}$ are

$$v_i^i(x) = \begin{cases} 0, & x = x_1, \\ -4, & x = x_2, \end{cases}$$

and for $i \neq j$,

$$v_j^i(x) = \begin{cases} c/\left(1-\delta_j\right), & x = x_1, \\ c/(1-\delta_j) - 4, & x = x_2. \end{cases}$$

We get $T_j(y, x, v_j^i)$ as in the tables below, where each cell contains the pair $(T_1(y, x, v_1^1), T_2(y, x, v_2^1))$. Due to symmetry we do not present the corresponding tables for $(T_1(y, x, v_1^2), T_2(y, x, v_2^2))$.

x_1	b_1	b_2
a_1	$\left(4, 4+\frac{c\delta_2}{1-\delta_2}\right)$	$\left(0, \frac{c}{1-\delta_2}\right)$
a_2	$\left(c, \frac{c\delta_2}{1-\delta_2}\right)$	$\left(1-4\delta_1, 1-\frac{c\delta_2}{1-\delta_2}-4\delta_2\right)$

x_2	b_1	b_2
a_1	$\left(0, \frac{c\delta_2}{1-\delta_2}\right)$	$\left(-4, -4+\frac{c}{1-\delta_2}\right)$
a_2	$\left(c-4, -4+\frac{c\delta_2}{1-\delta_2}\right)$	$\left(-3-4\delta_1, -3-\frac{c\delta_2}{1-\delta_2}-4\delta_2\right)$

The next step is to show that (a_1, b_2) is incentive compatible at both states for suitable δ_i, $i = 1, 2$ and c. Let us first consider state x_1. For the first player (column player) incentive compatibility is obtained when $1 - 4\delta_1 \leq 0$, i.e., $\delta_1 \geq 1/4$. Respectively, the second player's incentive compatibility condition gives us $c \geq 4(1 - \delta_2)$. These are also the conditions that we get for the incentive compatibility of (a_1, b_2) at state x_2. Because we can make a similar deduction for the second player we require

$$\min_{i\in I} \delta_i \geq 1/4 \text{ and } c \geq \max_{i\in I} 4(1 - \delta_i). \tag{9}$$

Even if the other action profiles (a_1, b_1) or (a_2, b_1) were incentive compatible, they cannot give the least payoff to $T_1\left(y, x, v_1^1\right)$ at any state. The action profile (a_2, b_2) on the other hand is not incentive compatible when $\delta_1 > 1/4$. For $\delta_1 = 1/4$ the outcomes (a_2, b_2) and (a_1, b_2) exactly the same for both players. This deduction gives

$$T_1\left(\mu^1(x), x, v_1^1\right) = \min_{y\in IC_1(v)(x)} T_1\left(y, x, v_1^1\right),\ x \in \{x_1, x_2\}.$$

Due to symmetry, we get the corresponding fixed-point property also for $T_2\left(\mu^2(x), x, v_2^2\right)$, when the parameters satisfy (9).

Similarly as above, we can observe that the stationary penal code $\{\mu^1, \mu^2\}$ with

$$\mu^1(x_1) = \mu^2(x_1) = (a_2, b_2),\ \mu^1(x_2) = (a_1, b_2),\ \mu^2(x_2) = (a_2, b_1),$$

satisfies (7) for $\delta_1, \delta_2 \leq 1/4$. Recall from earlier discussion that for this range of discount factors μ^1 and μ^2 are not Markov equilibria.

Assuming $c \geq \max_i 4(1 - \delta_i)$ we have candidates for an ESPC for any choice of discount factors. To verify that these candidates really are extremal it is sufficient to check that μ^1 and μ^2 give the players' security levels.

This example demonstrates that Markov strategies may give rise to ESPCs. Namely, we can take the policy profile $\mu^0(x_i) = (a_1, b_1)$, $i = 1, 2$ and ask whether the simply strategy $\{\mu^0, \mu^0\}$ is an ESPC. Let us assume $c \leq 4$ which guarantees that μ^0 is a Markov equilibrium. It is rather easy to see that then $\{\mu^0, \mu^0\}$ satisfies (7) for $v_i^j(x_1) = 4$ and $v_i^j(x_2) = 4\delta_i$, $i \in I$. Hence, for

$$\max_i 4(1 - \delta_i) \leq c \leq 4$$

we have at least two stationary penal codes that satisfy (7). However, $\{\mu^0 \mu^0\}$ is not extremal, i.e., (7) is only a necessary condition for an ESPC, and condition (8) is required to guarantee sufficiency.

In the next example we observe a corner solution property: the ESPCs are given by policies in which the players choose the smallest or largest y_i's among their feasible actions. In Section 5 we show that in deterministic resource extraction games the extremal payoffs are obtained when the players choose their largest efforts. In the following example we have the converse of this situation; the extremal payoffs are given by the minimal efforts.

Example 3. This example is a dynamic game of searching. The model originates from Diamond (1982) and has previously been studied by Curtat (1996), Horst (2005), and Nowak (2007). The state variable represents the productivity of search, $X = [1, 2]$, $W = X$. The actions are the efforts exerted by the players, $Y_i(x) = [0, 1]$ for all $x \in X$. The payoffs are

$$u_i(y, x) = xy_i \sum_{j \neq i} y_j - c(y_i),$$

where c is an increasing cost function. The transition is $f(y, x, w) = w$ where w follows a probability distribution with

$$p(\mathrm{d}w; y, x) = \frac{x + \sum_i y_i}{n + 2} g(w)\mathrm{d}w + \frac{n + 2 - x - \sum_i y_i}{n + 2} h(w)\mathrm{d}w.$$

The distribution defined by the density function g is assumed to stochastically dominate the distribution defined by h. In other words, high efforts are more likely to increase the productivity of search. As observed by Nowak (2007), the policy profile $\mu^m(x) = (0, \ldots, 0)$ is a Markov equilibrium of the game. Corresponding to this policy profile the players receive payoffs $v_j^i(x) = -c(0)/(1 - \delta_j)$, $i, j \in I$. We can see that $\{\mu^m, \ldots, \mu^m\}$ is an ESPC. Namely, even if $IC_i(v)(x)$ has other actions than 0, choosing $y_{-i} = (0, \ldots, 0)$ minimizes player i's payoff. Because μ^m is a Markov equilibrium, player i's best response is then to choose $y_i = 0$. Note also that μ^m gives the security level for each player. Hence, by Corollary 1 the minimum effort policy defines an extremal penal code.

4 Semi-Stationary Strategies

This section describes the class of equilibrium strategies that are supported by stationary simple strategies. This class is called semi-stationary equilibria. If the players follow a semi-stationary strategy profile σ, the resulting play is stationary, i.e., the future actions will depend only on the state after all histories when the players follow σ. However, if one of the player deviates from σ, i.e., from the ongoing policy profile, the rest of the game may be defined by another stationary policy profile. It is, however, assumed that deviations and the states at which they take place do not affect the future behavior.

Semi-stationary strategies are defined as follows.

Definition 4. A strategy profile σ is a semi-stationary strategy if

(1) after any history the players follow a stationary policy profile until some of them deviates unilaterally,
(2) if player i was the last to deviate unilaterally, the rest of the play is prescribed by a stationary policy profile that may only depend on i, the policy from which the player deviated, and the current stage.

A strategy is a semi-stationary equilibrium if it is subgame perfect and semi-stationary. These equilibrium strategies are denoted as Σ and the corresponding continuation payoff functions as $V(\Sigma)$. If after some history semi-stationary strategy prescribes a policy profile μ then the strategy is said to induce this policy profile. The induced policy profiles of semi-stationary equilibria are denoted as $M(\Sigma) \subseteq M$. Stationary Markov strategies as well as stationary simple strategies are semi-stationary strategies. In a repeated game, trigger strategies are semi-stationary strategies. However, we also have other semi-stationary strategies than trigger strategies.

The following example, which continues Example 3, shows one possible semi-stationary strategy in a search game, and presents the policy profiles that it induces.

Example 4. Let us assume that there are two players in the search game of Example 3 and $c(y_i) = y_i$. One possible semi-stationary strategy can be defined as follows:

(1) First, the players follow the policy profile $\mu^0(x) = (1, 1)$.
(2) If either of the players has unilaterally deviated from μ^0, they play according to $\mu^1(x) = (1/x, 1/x)$ for all $x \in X$.
(3) If either of the player has first unilaterally deviated from μ^0 and then from μ^1, they play according to $\mu^m = (0, 0)$.

This strategy induces three policy profiles: μ^0, μ^1, and μ^m. It is worth noticing that μ^0, μ^1, and μ^m are all Nash equilibria of the one-shot game in which x is fixed. Furthermore, the strategy is not a stationary simple strategy. If the policy in (2) was replaced by an assumption that the strategy leads to μ^1 if $\bar{x} \leq x^k$ and otherwise to μ^2, this strategy would not be semi-stationary due to its state dependence.

The following result tells that any induced policy profile of a semi-stationary equilibrium strategy can be supported with a subgame perfect stationary simple strategy.

Proposition 3. *Under (A1)–(A3), $\mu \in M(\Sigma)$ if and only if for all there are $\mu^i \in M$ such that $\sigma(\mu, \mu^1, \ldots, \mu^n) \in \Sigma^s$.*

Proof. If $\mu \in M(\Sigma)$, then after any in which the players have followed μ, there are μ^i, $i \in I$, which are the policy profiles corresponding to unilateral deviations from μ, and the same semi-stationary strategy that induces μ also induces these policy profiles after unilateral deviations. Here we need the assumption that the future policies do not depend on the current or future states. This is guaranteed by the second assumption in the definition of a semi-stationary strategy.

Due to subgame perfection there is no profitable one-shot deviation from μ when the continuations for these deviations are given by μ^i, $i \in I$. Hence, for all $k \geq 0$ and h^k that is obtained in k stages when μ is followed we have

$$v_i(\mu) \geq \sup_{y_i \in Y_i\left(x(h^k)\right)} T_i\left(y_i, \mu_{-i}\left(x(h^k)\right), x(h^k), v_i(\mu^i)\right). \tag{10}$$

This means that after any history that is obtained by playing μ there is a stationary simple strategy profile that supports it, namely that of $\sigma(\mu, \mu^1, \ldots, \mu^n)$.

The other direction follows directly from the fact that stationary simple strategies are semi-stationary; $\sigma(\mu, \mu^1, \ldots, \mu^n) \in \Sigma^s$ implies that $\mu \in M(\Sigma)$. ☐

The following result is an immediate corollary of propositions 1 and 3.

Corollary 2. *There are semi-stationary equilibria if and only if there are a stationary penal code and payoff functions that satisfy (4).*

Another consequence is that when ESPCs are attained, then $\mu \in M(\Sigma)$ if and only if it can be supported with an ESPC. The result follows from Proposition 3 because, as observed in Remark 2, under (A4) all stationary simple equilibria can be supported with ESPCs.

Corollary 3. *Under (A1)–(A4), $\mu \in M(\Sigma)$ if and only if there is $\{\mu^1, \ldots, \mu^n\} \in$ ext(Σ^p) such that $\sigma(\mu, \mu^1, \ldots, \mu^n) \in \Sigma^s$.*

Corollary 3 is particularly important because it tells that semi-stationary equilibria are the equilibrium strategies among which ESPCs lead to extremal payoffs. Consequently, ESPCs can be used in finding the policy profiles that give the highest payoff to a player. If such a policy profile is found, it gives the highest payoff to a player in the class of semi-stationary strategies. Let us briefly sketch how these maximal policies can be found. Let us suppose that we want to find the policy profile μ that maximizes player i's discounted profits such that $\sigma(\mu, \mu^1, \ldots, \mu^n)$ is subgame perfect. Supposing that the least extremal payoffs are given by continuation payoff function $\bar{v}$ we can formulate the problem as follows. Find $\mu \in M$ such that

$$v_j(\mu) = T_j\left(\mu(x), x, v_j(\mu)\right) \ \forall j \in I, \tag{11}$$

$$\begin{aligned} T_i\left(\mu(x), x, v_i(\mu)\right) = \sup \Big\{ & T_i\left(y, x, v_i(\mu)\right) : \ y \in Y(x), \\ & T_j\left(y, x, v_j(\mu)\right) \geq \sup_{y_j \in Y_j(x)} T_j\left(y_j, y_{-j}, x, \bar{v}_j^j\right), \ \forall j \in I \Big\}. \end{aligned} \tag{12}$$

Observe that (12) means that $v_i(\mu)$ satisfies the usual Bellman equation when player i's payoffs are maximized subject to incentive compatibility constraints.

The below result follows directly from Corollary 3.

Remark 3. Under (A1)–(A4), $\mu \in M(\Sigma)$ is maximal for player i, i.e., $v_i(\mu') \leq v_i(\mu)$ for all $\mu' \in M(\Sigma)$, if and only if there is $\{\mu^1, \ldots, \mu^n\} \in \text{ext}(\Sigma^p)$ such that $v_j(\mu)$, $j \neq i$, satisfy (11) and $v_i(\mu)$ satisfies (12), and for all policy profiles $\mu' \in M$ for which these conditions hold $v_i(\mu') \leq v_i(\mu)$.

The requirement $v(\mu') \leq v(\mu)$ for all $\mu' \in M$ that satisfy (11) and (12) is needed because there is no guarantee of uniqueness. In particular, there may be several ESPCs and different ESPCs give different maximizers for player i's payoff.

5 Resource Extraction Games

5.1 Maximum Effort Policy as an ESPC

In this section the purpose is to apply the previous results to the class of games in which the state variable represents a resource stock. The state is one dimensional; $x \in X = \mathbb{R}_+$, and the actions are the harvesting efforts

$y_i \in Y_i = [0, a_i]$, $a_i \leq 1$, $i \in I$. The total harvest is

$$h(y, x) = \min \left\{ \sum_{i \in I} y_i, 1 \right\} x,$$

when the stock is initially x. The harvest is shared in proportion to the harvesting efforts; the harvest of player i is

$$h_i(y, x) = \begin{cases} y_i h(y, x) / \sum_i y_i, & \text{if } \min_i y_i > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The growth of the stock is deterministic and governed by

$$f(y, x) = f(x - h(y, x)),$$

where f is increasing, $f(0) = 0$, and $f(s) > 0$ for all $s > 0$. The payoffs are assumed to be of the form $u_i(h_i(y, x))$, where u_i is strictly increasing and bounded from below. The vector of maximum efforts is denoted as $a^m = (a_1, \ldots, a_n)$. The following result tells that under quite loose conditions an ESPC is given by the policy $\mu^m(x) = a^m$ for all $x \in X$. Note that if μ^m defines an ESPC, then it is a Markov equilibrium.

Proposition 4. *The policy profile μ^m determines an ESPC in a resource extraction game when $(\sum_j a_j) - a_i \geq 1$ for all $i \in I$.*

Proof. Let $v^m(x)$ denote the continuation payoff of μ^m, i.e., $v^m : X \mapsto \mathbb{R}$. Since f and u_i, $i \in I$ are increasing so are the components of v^m. Let us set $v = (v^m, \ldots, v^m)$. We can first observe that $IC_i(v)(x) = \{a^m\}$ for all $i \in I$. Namely, the maximum of

$$\left[u_i \left(h_i(y, x) \right) + \delta v_i^m \left(f \left(x - h(y_i, a_{-i}^m, x) \right) \right) \right]$$

over $y_i \in Y_i$ is obtained at $y_i = a_i$ because $(\sum_j a_j) - a_i \geq 1$ for all $i \in N$, which means that h is the same for all y_i when $y_{-i} = a_{-i}^m$. Because u_i is strictly increasing, the unique optimum is then a_i. Thus, we have $IC_i(v)(x) = \{a^m\}$. Because μ^m gives each player the security level, it follows from Corollary 1 that $\mu^i(x) = a^m$, $i \in I$, form an ESPC. $\square$

5.2 Example

This example is a variation of “the great fish war game” of Levhari and Mirman (1980), see also Cave (1987). The variable x represents a fish stock, and there are two countries harvesting the stock. Each country decides on its harvesting effort $y_i \in Y_i = [0,1]$. The dynamics of the stock is governed by

$$f(y,x) = [x - h(y,x)]^{\alpha},$$

$\alpha \in (0,1)$. In the original paper by Levhari and Mirman, the payoffs are logarithmic. In our formulation this would mean

$$u_i(h_i(y,x)) = \ln\left(y_i h(y,x)/\left(y_1 + y_2\right)\right),\ i = 1,2.$$

It is rather easy to observe that with this utility function the payoffs are not bounded from below, i.e., they go to minus infinity as the efforts go to one. This is because the stock will stay at zero level after it is harvested into extinction. When players can be threatened with arbitrarily small payoffs after deviations, then any strategy that gives finite payoffs is an equilibrium.

If the feasible actions were $X_i = [0,1)$, $i = 1,2,$, then there would be no ESPCs even if u_i, $i = 1,2$, were bounded from below, i.e., (A4) fails. In this case we would, however, have ε-subgame perfect penal codes. Moreover, it is possible that there are equilibria in stationary simple strategies even in this case. For example, Levhari and Mirman (1980) construct a Markov equilibrium μ such that $\mu_i(x) \in (0,1)$, $i = 1,2$.

To make this example more tractable let us assume that payoffs are increasing and symmetric, i.e., $u_i = u$ for $i = 1,2$. Moreover, we may set $u(0) = 0$, which implies that payoffs are bounded from below. Observe that by the definition of h, the payoffs are $u(0)$ when y_1 and y_2 go to zero. By Proposition 4, the maximum effort equilibrium gives the extremal payoffs. The corresponding payoffs are $v_j^i(x) = u_j(h_i(\mu^i(x),x)) = u(x/2)$, $j,i \in \{1,2\}$. Using these functions we can find the maximum payoff that a country can get with semi-stationary strategies, i.e., solution to (11) and (12). The maximum payoff on the right hand side of incentive compatibility condition of (12) is obtained by using the maximum effort, i.e., the highest payoff that player $-i$ can get by deviating is $u(x/(1+y_i))$.

Figure 1 illustrates approximate functions $v_1(x)$ and $v_2(x)$ when the first player's payoffs are maximized and u is the square root function, $X = [0,2]$, $\delta = 4/5$, $\alpha = 1/2$, and $Y_1 = Y_2 = [0,1]$. The approximations were obtained by solving (11) and (12) for a discretized model where the state and actions spaces are $\tilde{X} = \{0, 2\Delta, 4\Delta, \ldots, 2\}$, $\tilde{Y}_1 = \tilde{Y}_2 = \{0, \Delta, 2\Delta, \ldots, 1\}$, and $\Delta = 1/25$. Furthermore, the state transitions of the discretized model were obtained by rounding $f(y,x)$ to nearest grid point for all $y \in \tilde{Y}$ and $x \in \tilde{X}$. As can be observed, the payoff for the second player, $v_2(x)$, is close to the lower bound $u(x/2)$, as supposed to.

The solution of (11) and (12) for the discretized model was found numerically

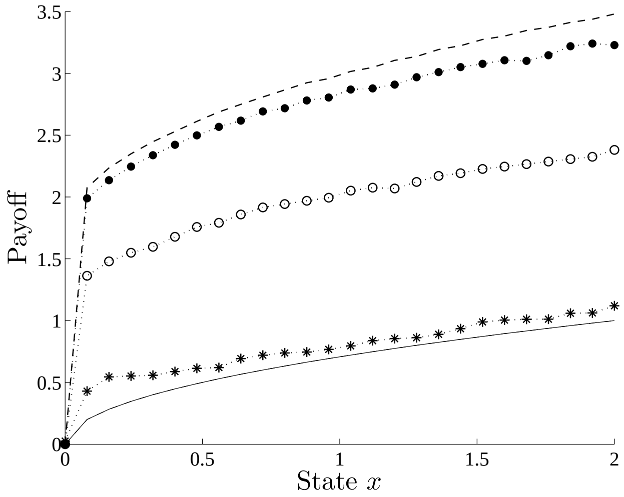

Figure 1. Payoffs of the first (•) and second (∗) player, when the first players payoff is maximized, the largest payoff to the first player among symmetric Markov equilibria (∘), the payoff at the ESPC (solid curve), and the payoff corresponding to the maximum of the first players profits over all policies (dashed curve).

by minimizing $\|v - Bv\|^2$ for an appropriately defined B with $v_i(x) = 2u(x/2)$, $i = 1, 2$, as initial values for $x \in \tilde{X}$. Instead of defining B as the right hand side of (11) and (12) for a given v, it was defined by taking the policy that gives the maximum in (12) and then evaluating the corresponding payoffs. This is analogous to the idea of policy iteration in dynamic programming. It was observed in numerical experiments note reported in this paper that this method is considerably more efficient the value iteration.

All the first player's payoff functions that can be attained with semi-stationary equilibria go between the minimal payoff (dashed curve) and the maximal one. Moreover, payoffs that can be obtained with symmetric Markov equilibria go between the function illustrated with circles and the dashed curve. As can be seen in Figure 1, the maximal payoff for the first player is considerably higher with semi-stationary strategies than with symmetric Markov strategies.

6 Conclusions

This paper has introduced semi-stationary and stationary simple strategies for dynamic games. The latter consist of policies that the players first follow and a set of policies corresponding to unilateral deviations of each player from the ongoing path. The policies that are followed after deviations form a penal code. The main result is that those penal codes that yield the minimum payoffs to players are characterized by a Bellman-type fixed-point equation. As observed in the paper, in some cases the extremal stationary penal codes are given by the players' security (min-max) strategies.

The extremal penal codes are of particular interest since they can be used in supporting any equilibria in the class of semi-stationary strategies. These are strategies that lead to behavior determined by stationary policies after all histories. Extremal stationary penal codes can be used in finding the highest payoffs that players can get in the set of semi-stationary strategies. An application to resource extraction was discussed. Under quite loose monotonicity assumptions, extremal penal codes are determined by players' maximum efforts at each state in these games.

The necessary condition for extremal stationary penal codes in the form of a fixed-point equation makes it possible to find them numerically by using methods for solving functional equations, or in the case of discrete state space, systems of equations. Compared to stationary Markov strategies this is a clear benefit. Because Markov strategies lack the structure that semi-stationary strategies have, finding extremal Markov equilibria is hard and may require finding all of them.

References

Abreu, D., 1986. Extremal equilibria of oligopolistic supergames. Journal of Economic Theory 39 (2), 191–225.

Abreu, D., 1988. On the theory of infinitely repeated games with discounting. Econometrica 56 (2), 383–396.

Benhabib, J., Radner, R., 1992. The joint exploitation of a productive asset: a game-theoretic approach. Economic Theory 2 (2), 155–190.

Bertsekas, D. P., Shreve, S. E., 1996. Stochastic Optimal Control: The Discrete Time Case. Athena Scientific, Belmont, Massachusetts.

Cave, J., 1987. Long-term competition in a dynamic game: The cold fish war. The RAND Journal of Economics 18 (4), 596–610.

Curtat, L., 1996. Markov equilibria of stochastic games with complementarities. Games and Economic Behavior 17 (2), 177–199.

Diamond, P., 1982. Aggregate demand management in search equilibrium. Journal of Political Economy 90, 881–894.

Duffie, D., Geanakoplos, J., Mas-Colell, A., McLennan, A., 1994. Stationary Markov equilibria. Econometrica 62, 745–781.

Horst, U., 2005. Stationary equilibria in discounted stochastic games with weakly interacting players. Games and Economic Behavior 51 (1), 83–108.

Levhari, D., Mirman, L. J., 1980. The great fish war: An example using a dynamic Cournot-Nash solution. The Bell Journal of Economics 11 (1), 322–334.

Mason, C. F., Polasky, S., Tarui, N., 2009. Cooperation on climate change mitigation. Working paper.

Nowak, A., 2007. On stochastic games in economics. Mathematical Methods of Operations Research 6, 513–530.

Nowak, A., Raghavan, T. E. S., 1992. Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. Mathematics of Operations Research 17, 519–526.

Parthasarathy, T., Sinha, 1989. Existence of stationary equilibrium strategies in non-zero sum discounted stochastic games with uncountable state space and state-independent transitions. International Journal of Game Theory 18 (2), 189–194.

Polasky, S., Tarui, N., Ellis, G. M., Mason, C. F., 2006. Cooperation in the commons. Economic Theory 29 (1), 71–88.

Shapley, L., 1953. Stochastic games. Proceedings of the National Academy of Sciences of the USA 39, 1095–1100.