xx manuscript No.
(will be inserted by the editor)

Extremal Pure Strategies and Monotonicity in Repeated Games

Kimmo Berg

Received: date / Accepted: date

Abstract The recent development of computational methods in repeated games has made it possible to study the properties of subgame-perfect equilibria in more detail. This paper shows that the lowest equilibrium payoffs may increase in pure strategies when the players become more patient and this may cause the set of equilibrium paths to be non-monotonic. A numerical example is constructed such that a path is no longer equilibrium when the players' discount factors increase. This property can be more easily seen when the players have different time preferences, since in these games the punishment strategies may rely on the differences between the players' discount factors. A sufficient condition for the monotonicity of equilibrium paths is that the lowest equilibrium payoffs do not increase, i.e., the punishments should not become milder.

Keywords repeated games · minimum payoff · monotonicity · equilibrium path · unequal discount factors · subgame perfection

1 Introduction

Repeated games have been studied extensively but it is still an open problem how exactly the set of subgame-perfect equilibria behaves when the players' time preferences change. Intuitively, it should be easier to support equilibria when the players are more patient, since patience puts more weight to the future payoffs. This makes the deviations less profitable since the continuation payoff is always higher on the equilibrium path compared to the punishment

K. Berg
Department of Mathematics and Systems Analysis, Aalto University School of Science, P.O. Box 11100, FI-00076 Aalto, Finland
Tel.: +31-618182860
Fax: +358-9-8632048
E-mail: kimmo.berg@aalto.fi

payoff that the player receives if he deviates from the equilibrium strategy. This intuition may, however, lead to wrong conclusions as it ignores one important factor: the punishment payoff may increase when the players become more patient, which may change the player's incentives and produce a profitable deviation. This paper examines the punishment strategies and the monotonicity of equilibria with respect to the discount factor. Many papers have focused on the properties of the equilibrium payoffs but here we pay special attention to the sequences of actions, i.e., the equilibrium paths.

The theory of infinitely repeated games has been developed in Abreu (1988) and Abreu et al. (1986, 1990). These papers characterize the set of subgame-perfect equilibria and show that the monotonicity of payoffs is related to the convexity of the payoff set; e.g., the equilibrium payoffs are convex and monotone when the players can use a public correlating device. Recently, Yamamoto (2010) has shown that without correlated strategies the payoff set is not in general convex nor monotone no matter how patient the players are, which is in sharp contrast to the folk theorem Fudenberg and Maskin (1986). The non-monotonicity of payoffs is also observed in Mailath et al. (2002), see also Mailath and Samuelson (2006), where they show that the maximum payoff in a prisoner's dilemma is decreasing for a certain range of discount factors.

This property can be easily seen in Figure 1, which shows the equilibrium payoffs in a prisoner's dilemma for two different discount factor values: $\delta = 0.4$ is shown by the plus signs and $\delta = 0.45$ (more patient players) is given by the smaller dots. We can see that most of the payoff points move a little when the players become more patient. For example, the maximum equilibrium payoff of player 1 inside the circle is given by the path where player 1 defects in the first stage giving payoff $(4, 0)$ and after that the players keep on cooperating, which gives them payoff $(3, 3)$. The average payoff of player 1 decreases as discounting shifts weight from the first stage payoff of 4 to the later payoffs of 3. The monotonicity of payoffs is related to the discreteness of the payoff set and the properties of average discounted payoffs but not the players' incentives nor the punishment strategies.

This paper shows that the monotonicity of paths is a more robust property of equilibria. The equilibrium paths are monotone in the discount factor if the punishment payoffs do not increase. This holds, e.g., in any prisoner's dilemma where the punishment payoffs remain the same for all discount factors. The monotonicity means that the players may design a sequence of actions and if this designed path is an equilibrium for a given level of patience, then it is also an equilibrium when the players become more patient. Moreover, this implies that the set of equilibrium paths may only enlarge. However, it is shown in this paper that the monotonicity of paths does not hold in general. A numerical example is constructed so that the punishment payoff increases when the players become more patient. This result raises a little warning to what may happen in a class of games and emphasizes the importance of the punishment strategies. It should be noted that the non-monotonicity of paths is not totally new observation since Mailath and Samuelson (2006) notice it under imperfect monitoring in Section 7.2.2.

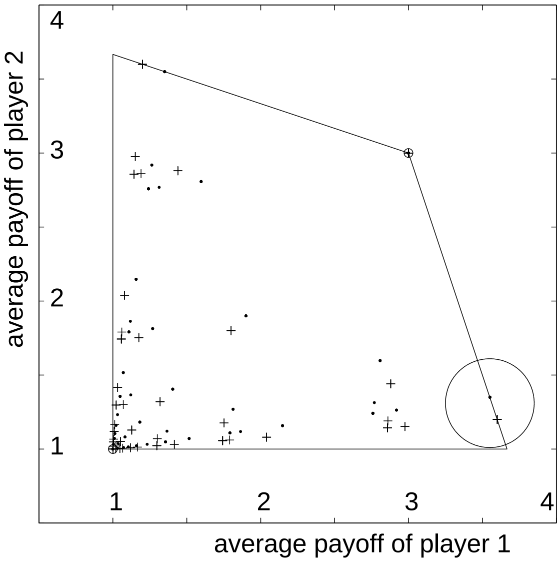

Fig. 1: Equilibrium payoffs in a prisoner's dilemma for two discount factors.

The relationship between the punishment payoffs and the stage game's minimax payoffs has received much attention in the literature, primarily related to the folk theorems (Fudenberg and Maskin, 1986, Abreu et al., 1994, Wen, 1994). The recent work has focused on the case of unequal discount factors (Lehrer and Pauzner, 1999, Salonen and Vartiainen, 2008, Houba and Wen, 2011, Guéron et al., 2011, Chen and Takahashi, 2012). Moreover, Gossner and Hörner (2010) examine the lowest equilibrium payoffs when the players cannot perfectly observe the opponents' actions. In these games, the lowest equilibrium payoff can be strictly lower than the minimax payoff; see Ex. 5.10 in Fudenberg and Tirole (1991). In this paper, we assume perfect monitoring and hence the punishment payoff is never below the minimax value.

Many computational methods for repeated games (Cronshaw and Luenberger, 1994, Cronshaw, 1997, Judd et al., 2003, Burkov and Chaib-draa, 2010, Salcedo and Sultanum, 2012, Abreu and Sannikov, 2014) are based on the set-valued fixed-point characterization of Abreu et al. (1986, 1990). Recently, Berg and Kitti (2012) have shown that the equilibrium paths consist of repeating fragments called elementary subpaths. This has provided a new methodology for analyzing the set of equilibria, computing the pure-strategy payoffs (Berg and Kitti, 2013) and identifying the equilibrium payoffs as particular fractals (Berg and Kitti, 2014). These papers assume that the punishment payoffs are known but it has turned out that the punishment paths may be very complicated and difficult to find in some games. This paper offers a simple solution for finding the punishment paths, when the set of equilibrium paths is small enough, i.e., when the discount factors are small. A better algorithm using the idea of branch and bound is presented in Berg and Kärki (2014a), but also it has problems finding the punishment paths in certain games. Solving this problem is left for future research and here we focus on examining the properties of the punishment paths.

The paper is structured as follows. In Section 2, the repeated game model is formulated and the notion of subgame-perfect equilibrium is defined. Section 3 examines the monotonicity of equilibrium paths and payoffs. A numerical method for finding the punishment paths is presented in Section 4. Three examples are given in Section 5; they demonstrate the punishment payoffs in an oligopoly game, the non-monotonicity of equilibrium paths, and the case of unequal discount factors. Section 6 concludes.

2 Subgame-Perfect Equilibria

The game has n players and $N = \{1, \dots, n\}$ denotes the set of players. The set of actions available to player i in the stage game is A_i. Each player is assumed to have finitely many actions. The set of action profiles is denoted by $A = \times_i A_i$. Moreover, a_{-i} denotes the action profile of other players than player i. The corresponding set of action profiles is $A_{-i} = \times_{j \neq i} A_j$. Function $u : A \mapsto \mathbb{R}^n$ gives the vector of payoffs that the players receive in the stage game when a given action profile is played, i.e., if $a \in A$ is played, player i receives payoff $u_i(a)$.

In the supergame, the stage game is repeated infinitely many times and the players discount the future payoffs with discount factors $\delta_i \in [0, 1)$, $i \in N$. Perfect monitoring is assumed: all players observe the chosen action profile at the end of each period. A history contains the path of action profiles that have been played before. The set of length k histories or paths is denoted by $A^k = \times_k A$. The empty path is $\varnothing$, i.e., $A^0 = \{\varnothing\}$, which corresponds to the history in the beginning of the game. Infinitely long paths are denoted by A^∞. When referring to the set of paths beginning with a given action profile a, notation $A^k(a)$ and $A^\infty(a)$ are used for length k and infinitely long paths, respectively. Moreover, $\mathcal{A}$ is the set of all paths, finite or infinite, and $\mathcal{A}(a)$ is the set of all paths starting with a, i.e., union of $A^k(a)$, $k = 1, 2, \dots$ and $A^\infty(a)$.

It is assumed that the players use pure strategies, i.e., randomized and correlated strategies are not allowed. A strategy for player i in the supergame is a sequence of mappings $\sigma_i^0, \sigma_i^1, \dots$, where $\sigma_i^k : A^k \mapsto A_i$. The set of strategies for player i is Σ_i. The strategy profile composed of $\sigma_1, \dots, \sigma_n$ is denoted by σ. Given a strategy profile σ and a path p, the restriction of the strategy profile after p is $\sigma|p$. The outcome path induced by σ is $(a^0(\sigma), a^1(\sigma), \dots) \in A^\infty$, where $a^k(\sigma) = \sigma^k(a^0(\sigma), \dots, a^{k-1}(\sigma))$ for all k.

The average discounted payoff of player i corresponding to a strategy profile σ is

$$U_i(\sigma) = (1 - \delta_i) \sum_{k=0}^{\infty} \delta_i^k u_i(a^k(\sigma)). \tag{1}$$

A strategy profile σ is a subgame-perfect equilibrium (SPE) of the supergame if

$$U_i(\sigma|p) \geq U_i(\sigma_i', \sigma_{-i}|p) \text{ for all } i \in N,\ p \in A^k,\ k \geq 0, \text{ and } \sigma_i' \in \Sigma_i.$$

From now on, equilibrium refers to subgame-perfect equilibrium. This paper focuses on SPE paths defined below.

Definition 1 A path $p \in A^\infty$ is a subgame-perfect equilibrium path if there is an SPE strategy profile that induces it.

It was shown in Abreu (1988), see also Abreu et al. (1986, 1990), that it is enough to study simple strategies when analyzing the set of equilibria. A simple strategy consists of an equilibrium path that is played and a punishment path for each player. The punishment paths are credible, i.e., they are the equilibrium paths that give the player's lowest equilibrium payoff. The players follow the current path unless there is a unilateral deviation by some player. In that case, the punishment path of the deviator becomes the path to be played. Again, if someone deviates from this new path, then this new deviator's punishment path becomes the one to be followed. If more than one player deviates from the current path, then the play remains on the given path. Since we examine non-cooperative games and the equilibrium notion is Nash equilibrium, the deviations by two or more players need not be considered. Thus, it is enough to check that no player alone should deviate when the deviation is followed by the punishment path of the player.

It should be noted that there are more complicated equilibrium strategies than simple strategies and some paths can be sustained with milder punishments, but Abreu's result shows that any equilibrium outcome, no matter how complicated, can also be implemented using simple strategies. Thus, there is no loss of generality in restricting attention to the simple strategies, or simply the paths that have no profitable unilateral deviations.

Let V be the compact set of SPE payoffs and $v_i^-(\delta)$ is the lowest SPE payoff of player i when the players have the discount factor $\delta = (\delta_1, \ldots, \delta_n)$, if V is non-empty. It should be noted that the set of equilibria may be empty in pure strategies. For this reason, we assume that the stage game has at least one Nash equilibrium in pure strategies, which implies that the set of equilibria is non-empty in the supergame. The equilibrium conditions for the SPE paths are given by the following one-shot deviation principle. A path $p = a^0(\sigma)a^1(\sigma)\cdots$ induced by a strategy σ is an SPE path if and only if

$$(1-\delta_i)u_i\left(a^k(\sigma)\right) + \delta_i v_i^k \geq \max_{a_i \in A_i}\left[(1-\delta_i)u_i\left(a_i, a_{-i}^k(\sigma)\right) + \delta_i v_i^-(\delta)\right], \quad (2)$$

for all $i \in N$, $k \geq 0$, and where

$$v_i^k = (1-\delta_i)\sum_{j=0}^{\infty}\delta_i^j u_i\left(a^{k+1+j}(\sigma)\right)$$

is the continuation payoff after $a^k(\sigma)$. The incentive compatibility (IC) condition (2) means that any player i should prefer the payoff given by path p

to any deviations at any stage that are followed by the punishment path with the payoff $v_i^-(\delta)$. Note that we can assume without loss of generality that the players deviate optimally and they receive the punishment payoffs after deviations. If an equilibrium can be sustained with a milder punishment then it can also be sustained with the optimal punishment. For example, the repetition of the stage game Nash equilibrium has no profitable deviations and therefore playing it does not require any punishment at all. Moreover, if a path has no optimal deviations then it does not have any deviations at all; this is why we have the maximum at the right-hand side of Eq. (2).

The definition of equilibrium is recursive in the sense that all equilibrium paths depend on the punishment payoffs $v^-(\delta)$ and the punishment paths depend on each other. In general, the payoffs $v^-(\delta)$ are not known, but with perfect monitoring they are above the minimax values, $v_i^-(\delta) \geq \underline{v}_i$, where

$$\underline{v}_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i}). \tag{3}$$

The aim of this paper is to find the punishment paths and the corresponding lowest equilibrium payoffs for different discount factors and analyze their properties.

3 Monotonicity of Equilibria

In this section, we examine the monotonicity of equilibrium paths and payoffs. Let us define the incentive compatibility with respect to a set of continuation payoffs W. Let $W \in \mathbb{R}^n$ be a non-empty, compact set and the punishment payoffs in the set are denoted by

$$v_i^-(W) = \min\{w_i,\ w \in W\}.$$

For player $i \in N$, the best possible deviation from an action profile $a \in A$ is to play the pure action that gives

$$d_i(a) = \max_{a_i \in A_i} u_i(a_i, a_{-i}).$$

Now, we say that a pair (a, w) of an action profile $a \in A$ and a continuation payoff $w \in W$ is admissible with respect to W if it satisfies the incentive compatibility conditions:

$$(1 - \delta_i)u_i(a) + \delta_i w_i \geq (1 - \delta_i)d_i(a) + \delta_i v_i^-(W),$$

for all $i \in N$. Let $C_a(W)$ denote the set of admissible continuation payoffs in W after action profile a:

$$C_a(W) = \{w \in W \text{ such that } (a, w) \text{ is admissible}\}.$$

Let us define a mapping $B : \mathbb{R}^n \mapsto \mathbb{R}^n$

$$B^\delta(W) = \bigcup_{(a,w) \in A \times W} (I - T)u(a) + Tw, \tag{4}$$

where (a, w) is admissible with respect to W, I is an $n \times n$ identity matrix, and T is a diagonal matrix with $\delta_1, \ldots, \delta_n$ on the diagonal.

Now, we are ready to present the fixed-point characterization of equilibrium payoffs (Abreu et al., 1986, 1990, Cronshaw and Luenberger, 1994, Mailath and Samuelson, 2006, Berg and Kitti, 2014).

Theorem 1 *The payoff set V is the largest fixed point of B:*

$$W = B^\delta(W) = \bigcup_{a \in A} \bigcup_{w \in C_a(W)} (I - T)u(a) + Tw.$$

It can also be shown that the payoff set V is compact. Note that the mapping B goes through all the action profiles $a \in A$ and all the possible continuations payoffs w that can follow a, i.e., the set $C_a(W)$. See Berg and Schoenmakers (2014) for the corresponding characterization in randomized strategies in a model where the players only observe the realized pure actions. There are many algorithms that use the iteration of B in computing approximations of the set of equilibrium payoffs (Cronshaw, 1997, Judd et al., 2003, Burkov and Chaib-draa, 2010, Salcedo and Sultanum, 2012, Abreu and Sannikov, 2014).

Let $V(\delta)$ denote the payoff set for a discount factor δ. It can be shown that the set of equilibrium payoffs is monotone in the discount factor when the set is convex (Abreu et al., 1990, Mailath and Samuelson, 2006).

Theorem 2 *Suppose $V(\delta_1)$ is convex, then $V(\delta_1) \subseteq V(\delta_2)$ for $\delta_2 \geq \delta_1$.*

Proof If W is self-generating for δ_2, i.e., $W \subseteq B^{\delta_2}(W)$, then $W \subseteq V(\delta_2)$ by Theorem 1 since $V(\delta_2)$ is the largest fixed-point of B^{δ_2}. Thus, it is enough to show that for all $v \in V(\delta_1)$ implies that $v \in B^{\delta_2}(V(\delta_1))$. Let $\delta_2 = \delta_1 + \epsilon$, where $\epsilon_i \geq 0$ for all $i \in N$.

Since $v \in V(\delta_1)$, there is an admissible pair (a, w^1) of an action profile $a \in A$ and a continuation payoff $w^1 \in V(\delta_1)$ such that $v = (I - T_1)u(a) + T_1 w^1$. We need to show that there is an admissible pair (a, w^2) with a continuation payoff $w^2 \in V(\delta_1)$ such that $v = (I - T_2)u(a) + T_2 w^2$. Note that the payoff v is between $u(a)$ and w^1 as a convex combination; and the same is true for $u(a)$ and w^2.

Now, we show that w^2 is between v and w^1, which implies that $w^2 \in V(\delta_1)$ since $v \in V(\delta_1)$, $w^1 \in V(\delta_1)$ and $V(\delta_1)$ is a convex set. We solve w^2 from the equations for v:

$$\begin{aligned}(I - T_1)u(a) + T_1 w^1 &= (I - T_2)u(a) + T_2 w^2 \\ \Rightarrow (T_1 + E)w^2 &= T_1 w^1 + \epsilon u(a),\end{aligned}$$

where E is a diagonal matrix with $\epsilon_1, \ldots, \epsilon_n$ on the diagonal. Thus, w^2 is a convex combination of w^1 and $u(a)$. Since v is between $u(a)$ and both w^1 and w^2, this means that w^2 is between w^1 and v.

Finally, let us check the admissibility of (a, w^2) with δ_2. Since the left-hand side of the incentive compatibility condition is the same for both pairs, it remains to show that the right-hand side, i.e., $v_i^-(V(\delta))$, is non-increasing in δ. Since $v_i^-(V(\delta_1)) \in V_i(\delta_1)$, the above implies that $v^-(V(\delta_2)) \leq v^-(V(\delta_1))$ and thus $V(\delta_1) \subseteq V(\delta_2)$. $\square$

It is easy to see that the payoff set may not be monotone when it is not convex, i.e., when the discount factors are small and the payoff set is discrete; see Fig. 1. The non-monotonicity is also demonstrated in Section 5.1, where the punishment payoffs increase when the players become more patient.

Theorem 2 implies that the payoff set is convex and monotone in the discount factor when the correlated strategies are available, which convexifies the payoff set (Abreu et al., 1990, Mailath and Samuelson, 2006). Let $co(W)$ denote the convex hull of set W and V^C is the payoff set with public correlation.

Proposition 1 *The payoff set V^C is the largest fixed point of $co(B)$:*

$$W = co\left(B^{\delta}(W)\right).$$

Thus, $V^C(\delta)$ is convex and monotone in δ.

Let us return back to pure strategies. The following shows that the convex, self-generating sets are monotone in the discount factor if the punishment payoffs do not increase.

Proposition 2 *Suppose $W \subseteq V(\delta_1)$ is convex, $W \subseteq B^{\delta_1}(W)$ and $v^-(V(\delta_1)) \geq v^-(V(\delta_2))$, then $W \subseteq V(\delta_2)$ for $\delta_2 \geq \delta_1$.*

Proof Again, it is enough to show that W is self-generating for δ_2, i.e., $W \subseteq B^{\delta_2}(W)$. The proof is similar to Theorem 2. The continuation payoff w^2 belongs to W since it is convex and (a, w^2) is admissible since the punishment payoff is non-increasing. □

When the discount factors are small, the payoff set consists of discrete number of points, as illustrated in Fig. 1, and the payoff set is neither convex nor monotone. Yamamoto (2010) presents a game where the payoff set is neither convex nor monotone no matter how large the discount factor is. This result suggest that there is a class of games that do not become dense for high levels of patience, and characterizing this class is an interesting research question. However, the set of equilibrium paths typically increases and the payoff set becomes more dense when the players become more patient. For example, Berg and Kärki (2014b) examine the lowest discount factor values when the payoff set covers all the reasonable payoffs in the symmetric 2×2 games under pure, randomized and correlated strategies, and this also gives a bound when the payoff set becomes convex and monotone. We give the result for the prisoner's dilemma; see Berg and Kärki (2014b), Berg and Schoenmakers (2014), Stahl (1991) for more details. Let $V^*(\delta) = \left\{v \in V^{\dagger} |\; v_i \geq v_i^-(V(\delta))\, \forall i \in N\right\}$ be the set consisting of the feasible and individually rational payoffs in the discounted game, where $V^{\dagger} = co\,(v \in \mathbb{R}^n : \exists a \in A \text{ s.t. } v = u(a))$ is the set of feasible payoffs in the undiscounted game.

Proposition 3 *In a symmetric prisoner's dilemma*

$\mathfrak{a}, \mathfrak{a}$ (a)	$\mathfrak{b}, \mathfrak{c}$ (b)
$\mathfrak{c}, \mathfrak{b}$ (c)	$\mathfrak{d}, \mathfrak{d}$ (d)

with $\mathfrak{c} > \mathfrak{a} > \mathfrak{d} > \mathfrak{b}$*, the payoff set* $V(\delta)$ *is convex and monotone in* δ *and covers all the feasible and individually rational payoffs when* $\delta \geq \delta^P$ *(*δ^M *or* δ^C*), where* P*,* M *and* C *refer to pure, randomized and correlated strategies:*

$$\delta^P = \delta^M = \frac{\mathfrak{c}-\mathfrak{b}}{\mathfrak{a}+\mathfrak{c}-\mathfrak{b}-\mathfrak{d}} > \max\left[\frac{\mathfrak{c}-\mathfrak{a}}{\mathfrak{c}-\mathfrak{d}}, \frac{\mathfrak{d}-\mathfrak{b}}{\mathfrak{a}-\mathfrak{b}}\right] = \delta^C,$$

when $\mathfrak{b}+\mathfrak{c} < 2\mathfrak{a}$*, and otherwise*

$$\delta^P = \frac{2(\mathfrak{c}-\mathfrak{d})}{\mathfrak{b}+3\mathfrak{c}-4\mathfrak{d}} > \delta^M = \frac{\mathfrak{c}-\mathfrak{b}}{2(\mathfrak{c}-\mathfrak{d})} > \frac{\mathfrak{d}-\mathfrak{b}}{\mathfrak{c}-\mathfrak{d}} = \delta^C.$$

Proof Let the letters a to d denote the action profiles, as shown by the game matrix; b refers to the action profile that gives the payoff $(\mathfrak{b}, \mathfrak{c})$. By Theorems 1 and 2, it is enough to find the smallest discount factor when $B^\delta(V^*(\delta))$ covers all the payoffs in $V^*(\delta)$. Note that the punishment payoff is a constant; $v_i^-(\delta) = \mathfrak{d}$ for all δ. In pure strategies, the last payoff to be filled by $B^\delta(V^*(\delta))$ depends on the shape of $V^*(\delta)$. In the quadrilateral shape, when $\mathfrak{b}+\mathfrak{c} < 2\mathfrak{a}$, the last point is on the efficient boundary between the payoffs $u(a)$ and $u(b)$ (or symmetrically $u(c)$). Thus, the required value of the discount factor, $\delta^P = \frac{\mathfrak{c}-\mathfrak{b}}{\mathfrak{a}+\mathfrak{c}-\mathfrak{b}-\mathfrak{d}}$, is solved when the sets $B_a^\delta(V^*(\delta))$ and $B_b^\delta(V^*(\delta))$ intersect, where $B_a^\delta(W) = \bigcup_{w \in C_a(W)} (I-T)u(a) + Tw$. Since these payoffs are obtained by playing pure strategies, the limit is the same in randomized strategies, i.e., $\delta^P = \delta^M$.

In the triangle shape, the last point to be filled in pure strategies is in the middle of the set $V^*(\delta)$ and the discount factor is solved when the sets B_b, B_c and B_d intersect. In randomized strategies, it is enough that the sets B_b and B_c intersect. This is a necessary condition since the efficient payoffs between $u(b)$ and $u(c)$ are only obtained by playing the pure strategies b and c, but it is also sufficient as then the payoffs inside the set $V^*(\delta)$ are filled. In correlated strategies, it is enough that the sets B_a and B_b cover the corner points of $V^*(\delta)$. The exact calculations are done in Berg and Kärki (2014b), Berg and Schoenmakers (2014).

Although, the convexity is typically required for the monotonicity of payoffs, it is not required for the monotonicity of paths, which is a more robust property of equilibria. A sufficient condition for the monotonicity of paths is that the punishment payoffs do not increase (Berg and Kitti, 2012). It is shown in Section 5.2 that the equilibrium paths may not be monotone when the punishment payoffs increase.

Theorem 3 *Suppose a path* $p \in A^\infty$ *is an SPE path for* δ_1 *and* $v^-(\delta_1) \geq v^-(\delta_2)$*, then* p *is an SPE path for* $\delta_2 \geq \delta_1$*.*

Proof Let u^k denote the payoffs in stage k on path p, i.e., $u^k = u(a^k)$ when the action profile a^k is played on stage k. Also, let $d^k = d(a^k)$ denote the deviation payoffs, and T_1 and T_2 are the diagonal matrices corresponding the discount factors δ_1 and δ_2, respectively. Now, we can rewrite the incentive compatibility conditions for δ_1:

$$(I - T_1)u^k + T_1 \left[(I - T_1) \sum_{j=0}^{\infty} T_1^j u^{k+j+1} \right] \geq (I - T_1)d^k + T_1 v^-(V(\delta_1)),$$

for all $k \geq 0$. Let us rearrange the equation and multiply from left by $(I - T_1)^{-1}$:

$$S_1^k \doteq u^k - d^k + T_1 \sum_{j=0}^{\infty} T_1^j \left(u^{k+j+1} - v^-(V(\delta_1)) \right) \geq \mathbf{0} \text{ for all } k = 0, 1, \ldots.$$

Similar expression can be derived for S_2^k with T_2, and the purpose of the proof is to show that $S_2^k \geq \mathbf{0}$ for all $k \geq 0$, which means that the incentive compatibility conditions hold for T_2 along the path p.

We can solve the recursion which S_i^k satisfies:

$$S_i^k = u^k - d^k + T_i \left(d^{k+1} - v^-(V(\delta_i)) + S_i^{k+1} \right), \text{ for all } k \geq 0 \text{ and } i = 1, 2.$$

Note that the first part $u^k - d^k$ is a vector with non-positive components, which implies that the components of the second part $d^{k+1} - v^-(V(\delta_1)) + S_1^{k+1}$, $k \geq 0$, must be non-negative, since $S_1^k \geq \mathbf{0}$ due to incentive compatibility.

Let $\delta_2 = \delta_1 + \epsilon$ and E is the diagonal matrix corresponding $\epsilon \geq \mathbf{0}$. We can simplify the expression for δ_2:

$$\begin{aligned} S_2^k &\geq u^k - d^k + (T_1 + E) \left(d^{k+1} - v^-(V(\delta_1)) + S_2^{k+1} \right) \\ &= S_1^k + E \left(d^{k+1} - v^-(V(\delta_1)) + S_1^{k+1} \right) + (T_1 + E) \left(S_2^{k+1} - S_1^{k+1} \right), \end{aligned}$$

where the first inequality follows from the fact that $v^-(V(\delta_1)) \geq v^-(V(\delta_2))$. Now, we can write

$$\begin{aligned} S_2^k - S_1^k &\geq E \left(d^{k+1} - v^-(V(\delta_1)) + S_1^{k+1} \right) + (T_1 + E) \left(S_2^{k+1} - S_1^{k+1} \right) \\ &\geq (T_1 + E) \left(S_2^{k+1} - S_1^{k+1} \right), \end{aligned}$$

where the inequality follows from the earlier observed non-negativity of $d^{k+1} - v^-(V(\delta_1)) + S_1^{k+1}$. Now, we can use this recursion:

$$S_2^k - S_1^k \geq E \sum_{j=0}^{\infty} (T_1 + E)^k Z_k,$$

where $Z_k = d^{k+1} - v^-(V(\delta_1)) + S_1^{k+1} \geq \mathbf{0}$. Thus, $S_2^k \geq S_1^k \geq \mathbf{0}$. □

Another useful property of the incentive compatibility conditions is that they are monotone in the punishment payoffs. This means that the set of IC paths with a higher punishment value is always a subset of the IC paths with a lower punishment value. This is an important result as it implies that the punishment paths and payoffs can be found with an iterative method. The idea is to compute the incentive compatible paths with the punishment payoffs that are lower than the actual ones, e.g., the minimax payoffs. These lower bound punishment payoffs can be increased systematically until the exact punishment payoffs are found, as the following results suggest.

Proposition 4 *Suppose a path p is incentive compatible with the punishment payoffs z^1, then p is incentive compatible with the punishment payoffs $z^2 \leq z^1$.*

Proposition 5 *Let W be the set of payoffs that the incentive compatible paths yield when the punishment payoffs are $z \leq v^-(\delta)$. Suppose that $v_i^-(W) > z_i$ for some $i \in N$, then $v_i^-(\delta) \geq v_i^-(W)$. Otherwise, we have $v^-(\delta) = z$.*

Proof Assume that $v_i^-(\delta) < v_i^-(W)$. This means that there is an incentive compatible path p_i^- with the punishment payoffs $v^-(\delta)$ which gives the payoff $v_i^-(\delta)$. Proposition 4 implies that p_i^- is incentive compatible with z and thus $v_i^-(W) \leq v_i^-(\delta)$, which is a contradiction. □

4 Computational Method

We present a simple method for finding the punishment payoffs based on Proposition 5. The method first sets the punishment values to the minimax values and increases these values until the optimal punishment payoffs are found. For each punishment value, the incentive compatible paths are computed and then the path that gives the player's smallest payoff is found for each player. If these smallest payoffs coincide with the current punishment values, then we have found the optimal punishment paths and payoffs. Otherwise, the current punishment values are too low and they can be increased up to the values found. This process is presented in Algorithm 1.

The Step 1 of Algorithm 1 relies on the fact that the IC paths can be computed for the given punishment values z_i. The method of Berg and Kitti (2012, 2013) can be used but it has problems when the number of IC paths is large, i.e., when the discount factors are high or the game is large. However, it should be noted that not all IC paths need to be computed but only the ones that potentially produce the lowest equilibrium payoffs. Especially, if some feasible punishment paths are found, then all paths that have payoffs above these values for all players can be neglected. These issues related to improving the algorithm are not discussed in this paper but rather left for future research. See Berg and Kärki (2014a) for a better algorithm using the branch and bound method. It should be noted that all the punishment paths found in this paper are optimal and computed with the presented method, unless otherwise indicated, since the numerical examples have finitely many

Algorithm 1: Find paths with lowest equilibrium payoffs

input : payoffs u_i, minimax values $\underline{v}_i$, discount factors δ_i.
output: paths p_i that give the punishment payoffs $v_i^-(\delta)$.

```
begin
    Initialize the lower bounds z_i = v_i.
    while punishment paths not found do
        1. Find incentive compatible paths with punishments z_i.
        2. Find the lowest payoffs m_i and paths p_i from the IC paths.
        if m_i > z_i for some i ∈ N then
            Update z_i = m_i.
        else
            Punishment paths and payoffs found.
```

elementary subpaths. If there are too many subpaths, we restrict their maximum length and in this case the method only produces feasible upper bounds for the punishment payoffs.

When the IC paths are found, the lowest payoffs for each player in Step 2 can be found in $O(mn)$ time (Madani et al., 2010, Papadimitriou and Tsitsiklis, 1987), where n is the number of nodes and m is the number of edges in the finite graph of IC paths. The task is essentially the same as finding the optimal strategies for discounted, infinite-horizon, deterministic Markov decision processes (DMDPs). Thus, this part of the algorithm can be solved efficiently.

This also means that the highest equilibrium payoffs can be found efficiently. Moreover, it is possible to find the maximum of any weighted sum of players' utilities with the method if the players have the same discount factors. There might be an efficient algorithm for the case of unequal discount factors, but the author is not aware of such method. The reason to solve the weighted sum problem is the fact that these solutions are Pareto efficient. However, there may be Pareto solutions that cannot be found as a weighted sum of players' utilities, and finding the whole Pareto frontier may be in general a difficult task.

5 Numerical Examples

The following examples demonstrate that the punishment payoff may increase when the players' discount factors increase. This fact is utilized in the second example such that the non-monotonicity of equilibrium paths is observed. The third example shows that this property can be more easily seen when the players have unequal discount factors.

5.1 Oligopoly Game

Let us examine the following quantity-setting duopoly game of Abreu (1988):

	L	M	H
L	$10, 10$ (a)	$3, 15$ (b)	$0, 7$ (c)
M	$15, 3$ (d)	$7, 7$ (e)	$-4, 5$ (f)
H	$7, 0$ (g)	$5, -4$ (h)	$-15, -15$ (i)

The firms may choose three output levels: low (L), medium (M) or high (H). The nine action profiles (L, L), (L, M), (L, H), $\ldots$ are denoted by letters a to i, and the stage game's Nash equilibrium is e, i.e., (M, M), giving payoff 7. The minimax payoff is $\underline{v}_i = 0$, and thus for all discount factors it holds that $0 \leq v_i^-(\delta) \leq 7$, $i = 1, 2$.

Let $(ab)^\infty = ababab\ldots$ denote that the path ab is played infinitely many times. Abreu (1988) examines the game when $\delta = 4/7$ and finds that the punishment paths are fb^∞ for player 1 and symmetrically hd^∞ for player 2. These paths give payoff -4 to the punished player in the first period and then payoff 3 in the following periods. The average discounted payoff is exactly the minimax value $v_i^-(\delta) = \underline{v}_i = 0$. Player 1 has no profitable one-shot deviations from the path fb^∞, since the deviation on the first period leads to path cfb^∞ giving (average discounted) payoff 0, and on the next periods the deviation leads to the path efb^∞ giving payoff 3, which is exactly the same as the continuation punishment path b^∞ with payoff 3. Also, player 2 should not deviate since the payoff $5(1 - \delta) + 15\delta \approx 10.7$ is higher than the payoff 3 of the deviation path ehd^∞. See the top part of Table 1 for the payoffs on the deviation and the continuation paths. Note also that the punishment paths are not unique with this discount factor and the paths c^∞ and g^∞ give the same zero payoff without any profitable deviations.

Table 1: The deviation and continuation payoffs for different paths.

path	player	stage	dev. path	dev. payoff	path	payoff
fb^∞	1	1	cfb^∞	0	fb^∞	0
fb^∞	1	2	efb^∞	3	b^∞	3
fb^∞	2	1	ehd^∞	3	fb^∞	$10\frac{5}{7}$
$f(db)^\infty$	1	1	$cf(db)^\infty$	$\frac{90741}{430000} \approx 0.21$	$f(db)^\infty$	$\frac{3129}{4300} \approx 0.73$
$f(db)^\infty$	1	3	$ef(db)^\infty$	$5\frac{77841}{430000} \approx 5.18$	$(bd)^\infty$	$5\frac{30}{43} \approx 5.70$
$f(db)^\infty$	2	1	$eh(bd)^\infty$	$5\frac{77841}{430000} \approx 5.18$	$f(db)^\infty$	$5\frac{87}{430} \approx 5.20$
$f(db)^\infty$	2	2	$eh(bd)^\infty$	$5\frac{77841}{430000} \approx 5.18$	$(db)^\infty$	$5\frac{30}{43} \approx 5.70$

The punishment payoffs for discount factors 0.25, 0.26, ..., 0.53 are shown in Figure 2. These can be computed with the method presented in Section 4. The punishment paths are optimal for $\delta \leq 0.38$ as the set of elementary subpaths is finite. For $\delta \geq 0.39$, the shown values are only upper bounds as the maximum length of the elementary subpaths had to be bounded for computational reasons; the subpaths were computed up to length 6. For these values, the method of Berg and Kärki (2014a) produces fast upper bound estimates in the range of 10^{-6}.

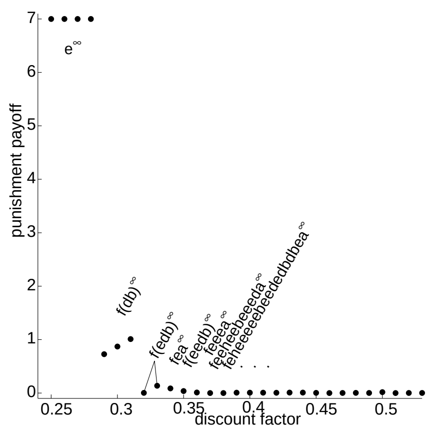

Fig. 2: The punishment payoffs for different discount factors.

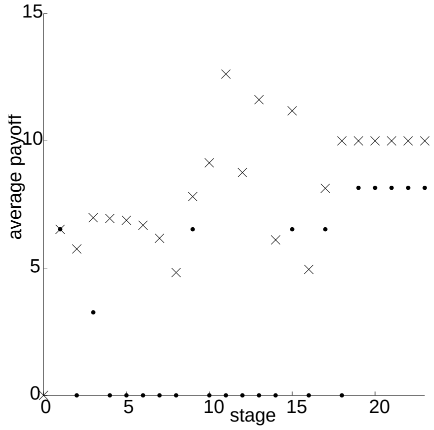

Fig. 3: The payoffs (x) and continuation payoff requirements $(\cdot)$ for $\delta = 0.38$.

When $\delta < 2/7$, the only SPE path is e^∞, i.e., the repetition of the Nash equilibrium. When $\delta = 2/7$, it is possible to play combinations of the elementary subpaths e, bd, db, fd and hb; all the equilibrium paths consists of these subpaths (Berg and Kitti, 2012, 2013). The punishment payoff for $\delta = 0.29$ is $3129/4300 \approx 0.728$ and it is given by the path $f(db)^\infty$ for player 1 and symmetrically $h(bd)^\infty$ for player 2. The bottom part of Table 1 shows that there are no one-shot deviations from these paths.

The punishment paths and all the SPE paths remain the same between the discount factors 0.29 and 0.31. However, the punishment payoff increases in this region, since the discounting puts more weight to the later payoffs 15 and 3 compared to the first period payoff -4. Thus, the punishment payoff is not monotone with respect to the discount factor.

In general, the path that gives the punishment payoff can be very simple or extremely complicated. When $\delta \geq 8/15$, it is possible to play c^∞ and the punishment payoff is then $v_i^-(\delta) = \underline{v}_i = 0$. On the other hand, the unique punishment path of player 1 is $feheeeeebeededbdbea^\infty$ with an approximate payoff $3 \cdot 10^{-10}$ when $\delta = 0.38$. Figure 3 shows the payoffs and the continuation payoff requirements of this path for each stage. For example, the first action profile f requires that the continuation payoff must be 6.53 and the continuation payoff given by the path $eheeeeebeededbdbea^\infty$ gives payoff $5 \cdot 10^{-10}$ above this value; we can see that the dot and the cross coincide at stage 1. On the other stages, the dots are below the crosses, which means that the incentive compatibility conditions are not binding and the continuation payoffs are well above the continuation payoff requirements.

In summary, we can note the following properties of the punishment paths and payoffs: 1) the payoff on the punishment path may go up and down when the punishment path is played; the only fact is that the first stage payoff is the smallest by definition, 2) the incentive compatibility condition is (almost) binding on the first round, i.e., the continuation payoff is close to the requirement, whereas there need to be no relation on the other rounds, and 3) the punishment path may be long and non-stationary. Actually, the last observation raises the question whether the players can determine and implement these long paths, or should some simplicity be assumed in the equilibrium behavior.

5.2 Non-monotonicity of Equilibrium Paths

Let us examine the following game:

	Q	W	E	R
Q	$5,5$ (a)	$3,4$ (b)	$-10,-10$	$-10,-10$
W	$4,3$ (c)	$2,2$ (d)	$-10,-10$	$-10,-10$
E	$-10,-10$	$-10,-10$	x,x (e)	$-10,y$ (f)
R	$-10,-10$	$-10,-10$	$y,-10$ (g)	$5,5$ (h)

where payoffs $y > x > 3 = \underline{v}_i$. Let us denote the action profiles (Q,Q), (Q,W), (W,Q) and (W,W) by a to d, and (E,E), (E,R), (R,E) and (R,R) by e to h. The first two columns and rows are set up such that it is possible to punish player 1 (2) by path b^∞ (c^∞) with payoff $v_i^-(\delta) = \underline{v}_i = 3$ if $\delta \geq 1/2$. However, the punishment payoff is above 3 and increases from $\delta = 0.34$ to $\delta = 0.35$. Now, values x and y are chosen such that the increase in the punishment payoffs makes a difference to the incentive compatibility of the other paths.

Let $x = 3.03$ and $y = 3.04$. The punishment payoff is approximately 3.0003 with path $(daaaeaea)^\infty$ when $\delta = 0.34$. With this punishment, it is possible to play e^∞ with payoff 3.03, since the deviation from e gives payoff $3.04(1-\delta)+3.0003\delta \approx 3.027 < 3.03$. Path e^∞ is not, however, an SPE path when $\delta = 0.35$. The punishment path is $(daaa)^\infty$ with approximate payoff 3.02 when $\delta = 0.35$. Now, there is a profitable deviation from e^∞, since $3.04(1-\delta)+3.02\delta \approx 3.033 > 3.03$. Thus, the set of equilibrium paths is not monotone with respect to the discount factor.

5.3 Unequal Discount Factors

Consider the following three-player game (Fudenberg and Maskin, 1986, Guéron et al., 2011, Chen and Takahashi, 2012):

C:

	L	R
T	$1,1,1$ (a)	$0,0,0$ (b)
B	$0,0,0$ (c)	$0,0,0$ (d)

D:

	L	R
T	$0,0,0$ (d)	$0,0,0$ (c)
B	$0,0,0$ (b)	$1,1,1$ (a)

which has essentially four action profiles: a with payoff 1 and $b/c/d$ with payoff 0 where player 2/1/3 may deviate to a. The minimax payoff is $\underline{v}_i = 0$, $i = 1,2,3$. In pure strategies, it is clear that a^∞ with payoff 1 is the only SPE outcome when the players have the same discount factor. This is because the NEU condition is violated and the punishers have no incentive to punish the deviators as they simultaneously punish themselves. On the other hand, it was shown by Guéron et al. (2011), Chen and Takahashi (2012) that the players can receive payoffs arbitrarily close to zero when the players have different discount factors.

Now, let us try to find some equilibria in pure strategies with a payoff less than one, when the players have different time preferences. The punishments rely on the differences in the discount factors and for example $\delta = (0.15, 0.65, 0.95)$ will serve our purpose. The path $d^{13}a^{\infty}$ for players 1 and 2 and the path $(ba)^{\infty}$ for player 3 are the optimal punishment paths. These give very close to zero payoffs to players 1 and 2 and approximate payoff 0.49 to player 3; see Table 2 for the payoffs. Players 1 and 2 cannot deviate from path $d^{13}a^{\infty}$ and player 3 has no incentive ($0.5133 > 0.5128$). Also, player 3 cannot deviate from path $(ba)^{\infty}$ and player 2 enjoys this path with payoff 0.394 compared to a deviation followed by the path $d^{13}a^{\infty}$ with payoff 0.352. This example shows that the players may receive payoffs less than one in this game when they have different discount factors, use pure strategies and have no public correlating devices. Thus, it shows that the equilibrium paths are not monotone because all the discount factors can be increased to the same value and then path a^{∞} is the only equilibrium.

Table 2: The players' approximate payoffs from different paths.

path	player 1	player 2	player 3
$d^{13}a^{\infty}$	$1.9 \cdot 10^{-11}$	$3.7 \cdot 10^{-3}$	**0.5133**
$a(ba)^{\infty}$	0.870	0.606	**0.5128**
$(ba)^{\infty}$	0.130	**0.394**	0.487
$ad^{13}a^{\infty}$	0.85	**0.352**	0.538

6 Conclusion

This paper examines the monotonicity of pure-strategy subgame-perfect equilibria with respect to the discount factor in infinitely repeated games with perfect monitoring. It is shown that the equilibrium paths are monotone if the punishment payoffs do not increase. The monotonicity means that a certain sequence of action profiles remain equilibrium when the players become more patient. Thus, the players do not need to alter their actions, but the change in the discount factors may affect their payoffs. This paper constructs an example where the punishment payoffs increase, which introduces a violation in the incentive compatibility conditions and this shows the non-monotonicity of equilibrium paths.

The monotonicity of paths is a more robust property and different from the monotonicity of payoffs, which has been observed in the literature. The monotonicity of payoffs is caused by the discreteness of the payoff set and the discounted payoff criterion. It may be that the set of equilibrium paths remains exactly the same but the payoff set is not monotone, as the payoffs may shift a little. The monotonicity of payoffs is guaranteed by the convexity of the payoff set, which may be difficult to satisfy. On the contrary, it is much easier to satisfy the condition that the punishment payoffs do not increase. For example, many of the 2×2 games have a constant punishment payoff, whereas the convexity is guaranteed only when the discount factors are high (Berg and Kärki, 2014b).

This paper emphasizes the importance of the punishment paths and payoffs. They support the whole set of equilibria and tell what happens if the players deviate from the path of play. This paper provides an algorithm for finding the punishment paths when the set of incentive compatible paths is not too large, i.e., when the discount factors are small enough. It is found that the punishment paths may be long, non-stationary and difficult to find for a class of games. It seems that finding the punishment paths corresponds to solving a discrete optimization problem over the pure-strategy paths and this means that there is no special characterization or equation from which the punishment paths could be solved.

One interesting idea for future research is to generalize the punishment paths to stochastic games. The concept of elementary subpaths has been extended to stochastic games in Berg (2012), but the algorithm for computing equilibria or punishment paths has not been implemented yet. Essentially, the difference is that the punishment paths are state dependent in stochastic games. Finally, an open question is how general the non-monotonicity is in economically meaningful games. It seems that it is rare and more of a mathematical curiosity.

Acknowledgements The author thanks the reviewers for the valuable suggestions for improving the paper, and acknowledges funding from Emil Aaltosen Säätiö through Post doc -pooli.

References

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

Abreu, D., Dutta, P. K., & Smith, L. (1994). The folk theorem for repeated games: A Neu condition. *Econometrica*, 62 (4), 393–948.

Abreu, D., Pearce, D., & Stacchetti, E. (1986). Optimal cartel equilibria with imperfect monitoring. *Journal of Economic Theory*, 39 (1), 251–269.

Abreu, D., Pearce, D., & Stacchetti, E. (1990). Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica*, 58 (5), 1041–1063.

Abreu, D., & Sannikov, Y. (2014). An algorithm for two-player games with perfect monitoring. *Theoretical Economics*, 9 (2), 313–338.

Berg, K. (2012). Characterization of equilibrium paths in discounted stochastic games. Working paper.

Berg, K., & Kitti, M. (2012). Equilibrium paths in discounted supergames. Working paper. `http://sal.aalto.fi/publications/pdf-files/mber09b.pdf`

Berg, K., & Kitti, M. (2014). Fractal geometry of equilibrium payoffs in discounted supergames. *Fractals*. to appear. `http://sal.aalto.fi/publications/pdf-files/mber12.pdf`

Berg, K., & Kitti, M. (2013). Computing equilibria in discounted 2×2 supergames. *Computational Economics*, 41, 71–78.

Berg, K., & Kärki, M. (2014a). An Algorithm for Finding the Minimal Pure-Strategy Subgame-Perfect Equilibrium Payoffs in Repeated Games. Working paper.

Berg, K., & Kärki, M. (2014b). How patient the players need to be to get all the relevant payoffs in the symmetric 2×2 supergames? Working paper.

Berg, K., & Schoenmakers, G. (2014). Construction of randomized subgame-perfect equilibria in repeated games. Working paper.

Burkov, A., & Chaib-draa, B. (2010). An Approximate Subgame-Perfect Equilibrium Computation Technique for Repeated Games. *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence*, 729–736.

Chen, B., & Takahashi, S. (2012). A folk theorem for repeated games with unequal discounting. *Games and Economic Behavior*, 76 (2), 571–581.

Cronshaw, M. B. (1997). Algorithms for finding repeated game equilibria. *Computational Economics*, 10, 139–168.

Cronshaw, M. B., & Luenberger, D. G. (1994). Strongly symmetric subgame perfect equilibria in infinitely repeated games with perfect monitoring. *Games and Economic Behavior*, 6, 220–237.

Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting and incomplete information. *Econometrica*, 54 (3), 533–554.

Fudenberg, D., & Tirole, J. (1991). *Game Theory*. The MIT Press, Cambridge, MA.

Gossner, O., & Hörner, J. (2010). When is the lowest equilibrium payoff in a repeated game equal to the minmax payoff? *Journal of Economic Theory*, 145 (1), 63–84.

Guéron, Y., Thibaut, L., & Thomas, C. D. (2011). On the folk theorem with one-dimensional payoffs and different discount factors. *Games and Economic Behavior*, 73 (1), 287–295.

Houba, H., & Wen, Q. (2011). Extreme equilibria in the negotiation model with different time preferences. *Games and Economic Behavior*, 73, 507–516.

Judd, K., Yeltekin, Ş., & Conklin, J. (2003). Computing supergame equilibria. *Econometrica*, 71, 1239–1254.

Lehrer, E., & Pauzner, A. (1999). Repeated games with differential time preferences. *Econometrica*, 67, 393–412.

Madani, O., Thorup, M., & Zwick, U. (2010). Discounted deterministic Markov decision processes and discounted all-pairs shortest paths. *ACM Transactions on Algorithms*, 6 (2), 33:1–33:25.

Mailath, G. J., Obara, I., & Sekiguchi, T. (2002). The maximum efficient equilibrium payoff in the repeated prisoners' dilemma. *Games and Economic Behavior*, 40, 99–122.

Mailath, G. J., & Samuelson, L. (2006). *Repeated games and reputations: long-run relationships*. Oxford University Press.

Papadimitriou, C., & Tsitsiklis, J. N. (1987). The complexity of Markov decision processes. *Mathematics of Operations Research*, 12 (3), 441–450.

Salcedo, B., & Sultanum, B. (2012). Computation of subgame-perfect equilibria of repeated games with perfect monitoring and public randomization. Working paper.

Salonen, H., & Vartiainen, H. (2008). Valuating payoff streams under unequal discount factors. *Economics Letters*, 99 (3), 595–598.

Stahl, D. O. (1991). The graph of prisoners' dilemma supergame payoffs as a function of the discount factor. *Games and Economic Behavior*, 3, 368–384.

Wen, Q. (1994). The "folk theorem" for repeated games with complete information. *Econometrica*, 62 (4), 949–954.

Yamamoto, Y. (2010). The use of public randomization in discounted repeated games. *International Journal of Game Theory*, 39, 431–443.