

Article

Construction of Subgame-Perfect Mixed-Strategy Equilibria in Repeated Games

Kimmo Berg ^{1,*}  and Gijs Schoenmakers ²

¹ Department of Mathematics and Systems Analysis, Aalto University School of Science, P.O. Box 11100, FI-00076 Aalto, Finland

² Department of Data Science and Knowledge Engineering, Maastricht University, P.O. Box 616, 6200MD Maastricht, The Netherlands; gm.schoenmakers@maastrichtuniversity.nl

* Correspondence: kimmo.berg@aalto.fi; Tel.: +358-40-7170-025

Received: 26 July 2017 ; Accepted: 18 October 2017 ; Published: 1 November 2017

Abstract: This paper examines how to construct subgame-perfect mixed-strategy equilibria in discounted repeated games with perfect monitoring. We introduce a relatively simple class of strategy profiles that are easy to compute and may give rise to a large set of equilibrium payoffs. These sets are called self-supporting sets, since the set itself provides the continuation payoffs that are required to support the equilibrium strategies. Moreover, the corresponding strategies are simple as the players face the same augmented game on each round but they play different mixed actions after each realized pure-action profile. We find that certain payoffs can be obtained in equilibrium with much lower discount factor values compared to pure strategies. The theory and the concepts are illustrated in 2×2 games.

Keywords: repeated game; mixed strategy; subgame perfection; payoff set

MSC: 91A20

JEL Classification: C73

1. Introduction

This paper examines how mixed actions can be used in obtaining subgame-perfect equilibria in discounted repeated games. In repeated games, the set of subgame-perfect equilibria can be defined recursively: a strategy profile is an equilibrium if certain equilibrium payoffs are available as continuation payoffs, and these continuation payoffs may be generated by means of other equilibrium strategy profiles. This construction has been presented for pure strategies in Abreu et al. [1,2], where they give a fixed-point characterization of the set of equilibrium payoffs (see also [3–12]).

The mixed-strategy model has been examined in [3,13–15], where it is shown that the folk theorem holds with and without public correlation and observable mixed actions. We examine a model where correlated strategies are not available and the players are not arbitrarily patient but have fixed discount factors (not necessarily all the same) between zero and one. Busch and Wen [16] examine a related model of negotiation with unobservable mixed actions. The model has also been generalized to imperfect monitoring [2,17,18], incomplete information [19,20], and stochastic games [21–23].

It is difficult to compute the set of subgame-perfect equilibria in repeated games. This is because the equilibrium strategies may depend recursively on each other. To find an equilibrium strategy, one needs to know the equilibrium strategies that produce the continuation payoffs that the strategy requires, and these continuation payoffs may be different after each realized pure-action profile. The only method that we are aware of is by Berg [24] that systematically enumerates all the required stage games that emerge in the repeated game. This method has only been implemented in specific

situations and it has not been implemented in general games. One difficulty in computing equilibria is finding the optimal punishment payoffs and strategies (see [25,26], where pure-strategy punishments are studied). This is an open problem, and we assume that the punishment payoffs are known. In this paper, we introduce a new concept called self-supporting sets, which give strategy profiles that are easy to compute and generate a large set of equilibrium payoffs. The idea is based on the concept of self-generating sets [3], which are sets that can be generated as equilibrium payoffs in the repeated game such that the continuation payoffs are chosen from the set itself. We simplify the concept by requiring that the whole set is generated by Nash equilibria of a single stage game. Our idea relies on the fact that mixed strategies can generate an uncountable set of payoffs in a single stage game when the players are indifferent between the actions¹; note, the idea of indifference is also used in belief-free equilibrium [27,28]. On the equilibrium path, players face the same augmented game (that takes into account the continuation payoffs) on each round, but they may play different mixed actions after each realized pure-action profile. Thus, our strategies on the equilibrium path are simple and they can be presented as a finite state strategy equilibrium [29]². This simplicity has also a practical advantage as boundedly rational players are not likely to play complicated strategies. Moreover, the self-supporting strategies are not stationary nor of one-period memory when the pure-action profiles are used as states. Our idea eliminates the complicated recursive dependency of the self-generating sets and the set of equilibria. The self-generating sets are dramatically more complicated since they may involve long sequences of dependencies of payoffs that are generated by different strategies that require continuation payoffs that are generated by other strategies, which require some new continuation payoffs and so on (see Example 1 in Section 3.2).

The self-supporting sets and corresponding strategies are useful in the sense that they do not rely on nor require any other equilibrium profiles or payoffs, except for the punishment profiles in the case that a player deviates from his strategy. Thus, if one can find such sets, they are automatically subsets of the equilibrium payoffs. We demonstrate the concept in a prisoner's dilemma, where a large set of equilibrium payoffs is obtained with surprisingly low discount factor values, and this shows a dramatic difference to the pure-strategy equilibria. See [3,10,11,27,30–33] for the earlier analysis of prisoner's dilemma. Moreover, we believe that large, self-supporting sets can be found in many games since they exist, at least, in the following 2×2 repeated games: prisoner's dilemma, stag hunt, chicken and no conflict games (see Figure 1).

It is interesting to compare the set of equilibria in pure, mixed and correlated strategies. It is not clear what the correct model is and it depends on the game situation in hand. Pure strategies may be founded, e.g., in public or governmental decision making where it is unlikely that the players would randomize between the alternatives. This may also be the case with boundedly rational players who may prefer simple strategies rather than playing different mixed strategies on each round the game is played; note, in some models, it is possible to purify mixed-strategy equilibria [34,35]. The mixed strategies can, however, produce higher equilibrium payoffs, as was shown in [24], or decrease the punishment payoffs, which may enlarge the set of equilibria. A higher payoff itself is a reason enough to play mixed strategies, and it is an open problem how to find all mixed-strategy equilibria. In this paper, we show that the interior payoffs in the prisoner's dilemma can be obtained in mixed strategies with much lower discount factor values compared to the pure strategies. Finally, we note that the correlated strategies are not reasonable in all game situations, since it requires a trusted third party that organizes the public randomization and sends the suitable signals to players so that they can coordinate their actions. We also note that there are many models of correlation in repeated games as

¹ Figure 1 (and also the example in Section 5) shows that the self-supporting sets may produce a large portion of the area of the feasible and individually rational payoffs in the game and thus even larger portion of the equilibrium payoffs.

² Only the strategy on the equilibrium path is guaranteed to have a finite presentation but we are not aware of any result that the punishment strategies have a finite presentation.

the players may receive signals on every round the game is played or only before the first round, and the players may correlate on both pure or mixed strategies.

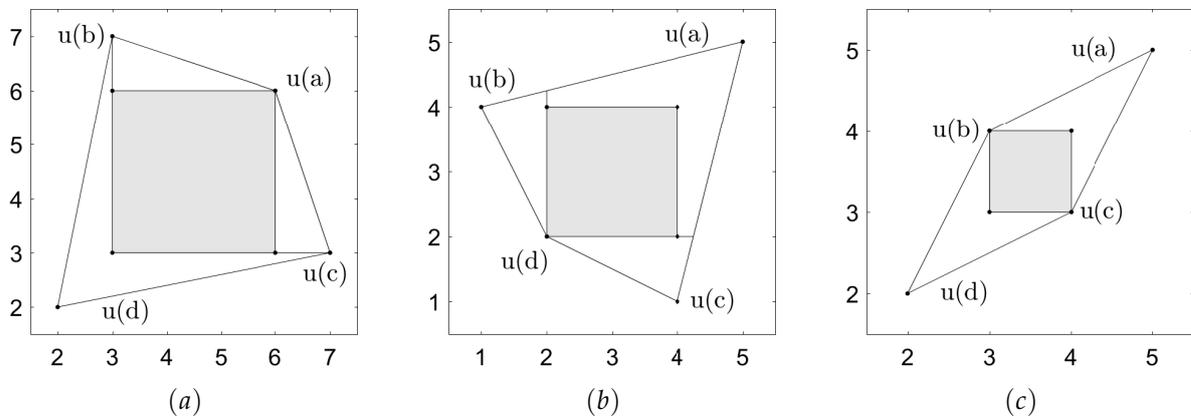


Figure 1. Self-supporting sets in (a) chicken, (b) stag hunt and (c) no conflict games. The players’ payoffs of action profiles a, b, c and d are denoted by $u(a)$ to $u(d)$, respectively.

This paper is structured as follows. Section 2 introduces stage games and the basics of 2×2 games. The repeated games are analyzed in Section 3. Section 4 examines a prisoner’s dilemma game. Section 5 demonstrates the self-supporting sets in a quantity-setting duopoly. Section 6 is the conclusion.

2. The Model

2.1. Stage Games

In a repeated game, a stage game is played repeatedly by the same players. A stage game is characterised by the tuple (N, A, u) , where

- $N = \{1, \dots, n\}$ is the finite set of players,
- A_i is the finite set of pure actions for player $i \in N$, and $A = \times_{i \in N} A_i$ is the set of pure-action profiles. Also, a pure action of player i is called $a_i \in A_i$ and a pure-action profile is called $a \in A$.
- $u \in \mathbb{R}^n$ is the payoff vector.

The play in a stage game: each player $i \in N$ is allowed to randomize over his pure actions $a_i \in A_i$, yielding a mixed action $q_i \in Q_i$, where Q_i is the set of probability distributions over A_i , and $Q = \times_{i \in N} Q_i$. So, it holds that $q_i(a_i) \geq 0$ for each $a_i \in A_i$ and $\sum_{a_i \in A_i} q_i(a_i) = 1$. A pure action is a mixed action as well. A mixed-action profile is denoted by $q = (q_1, \dots, q_n) \in Q$. The carrier (or support) of a mixed action is the set of pure actions that is played with a strictly positive probability: $Car(q_i) = \{a_i \in A_i | q_i(a_i) > 0\}$. We also define $Car(q) = \times_{i \in N} Car(q_i)$ and for each $a \in Car(q)$, we let $\pi_q(a)$ be the probability that the action profile a is realized if the mixed-action profile q is played: $\pi_q(a) = \prod_{j \in N} q_j(a_j)$. The payoffs in a stage game are given by the function $u : Q \mapsto \mathbb{R}^n$. If the players use a mixed-action profile $q \in Q$, then player i receives an expected payoff of³

$$u_i(q) = \sum_{a \in A} u_i(a) \pi_q(a). \tag{1}$$

³ Note that $u_i(a)$ denotes the stage-game payoffs and $u_i(q)$ the expected payoff of a mixed strategy q .

Player i 's opponents' action profile is denoted by $q_{-i} \in Q_{-i} = \times_{j \in N, j \neq i} Q_j$. An action profile q is a Nash equilibrium in the stage game if no player has a profitable deviation, i.e.,

$$u_i(q) \geq u_i(q'_i, q_{-i}) \text{ for all } i \in N, \text{ and } q'_i \in Q_i. \tag{2}$$

2.2. Equilibria in 2×2 Stage Games

In this subsection, we examine what the sets of equilibrium payoffs can be in 2×2 stage games. These games have many classifications [36]. For example, Borm [37] shows that there are 15 classes of games (denoted by c1–c15) based on the players' best responses. Here, we are interested in three types of games: (i) those that have a two-dimensional set of equilibria (typically Borm's class c1), (ii) one-dimensional set of equilibria (i.e., one or more line segments; classes c2–c4, c7, c10–c13), and (iii) zero-dimensional set of equilibria (i.e., only one or more points, namely, c5, c6, c9 and c14). We are mostly interested in the first two types of games, since they generate a larger set of payoffs. The following examples demonstrate each type of game.

Game 1 (c1)	Game 2 (c7)	Game 3 (c14)												
<table style="border-collapse: collapse; text-align: center;"> <tr><td style="border: 1px solid black; padding: 2px;">3, 3</td><td style="border: 1px solid black; padding: 2px;">1, 3</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">3, 1</td><td style="border: 1px solid black; padding: 2px;">1, 1</td></tr> </table>	3, 3	1, 3	3, 1	1, 1	<table style="border-collapse: collapse; text-align: center;"> <tr><td style="border: 1px solid black; padding: 2px;">7/3, 7/3</td><td style="border: 1px solid black; padding: 2px;">1/3, 3</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">11/3, 1</td><td style="border: 1px solid black; padding: 2px;">1, 1</td></tr> </table>	7/3, 7/3	1/3, 3	11/3, 1	1, 1	<table style="border-collapse: collapse; text-align: center;"> <tr><td style="border: 1px solid black; padding: 2px;">2, 1</td><td style="border: 1px solid black; padding: 2px;">0, 0</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">0, 0</td><td style="border: 1px solid black; padding: 2px;">1, 2</td></tr> </table>	2, 1	0, 0	0, 0	1, 2
3, 3	1, 3													
3, 1	1, 1													
7/3, 7/3	1/3, 3													
11/3, 1	1, 1													
2, 1	0, 0													
0, 0	1, 2													

Game 1 has a two-dimensional set of equilibrium payoffs, namely the convex hull of $(1, 1), (1, 3), (3, 1)$ and $(3, 3)$. Since both players are always indifferent, any mixed-action pair is an equilibrium. In Game 2, the top action is a strictly dominated action for player 1, and player 2 is indifferent between his actions when player 1 plays the bottom action. Thus, player 2 can play any mixed action when player 1 chooses the bottom action, leading to a one-dimensional set of equilibrium payoffs, namely the line segment $[1, 11/3] \times 1$. Game 3 is a battle-of-the-sexes game with three equilibria, two in pure actions and one in mixed actions. These generate three payoff points, where the mixed equilibrium gives the lowest payoff of $(2/3, 2/3)$.

3. Repeated Games

In a repeated game, a stage game is repeated infinitely often and we make the usual assumption that the players only observe the realized pure actions and not the accompanying mixed actions. This so-called public past play is denoted by the set of histories $H^k = A^k = \prod_k A$ for stage $k \geq 0$, where $H^0 = A^0 = \{\emptyset\}$ and corresponds to the beginning of the game. So, a history contains all the pure-action profiles that have been realized at the previous stages. The set of all histories is $H = \cup_{k=0}^{\infty} H^k$. In a repeated game, a public strategy σ_i of player $i \in N$ is a mapping that assigns a probability distribution over player i 's actions for each possible history $\sigma_i : H \mapsto Q_i$. The set of player i 's strategies is Σ_i . The players' strategies form the strategy profile $\sigma = (\sigma_1, \dots, \sigma_n)$, a strategy profile of all players except player i is denoted by σ_{-i} and the set of strategy profiles is given by $\Sigma = \times_{i \in N} \Sigma_i$.

Player i discounts the future payoffs with a discount factor $\delta_i \in (0, 1)$. The average discounted payoff of a strategy profile σ for player i is

$$U_i(\sigma) = (1 - \delta_i) \sum_{k=0}^{\infty} \delta_i^k u_i^k(\sigma), \tag{3}$$

where $u_i^k(\sigma)$ is the payoff of player i at stage k induced by the strategy profile σ . A profile σ is a Nash equilibrium if no player has a profitable deviation, i.e.,

$$U_i(\sigma) \geq U_i(\sigma'_i, \sigma_{-i}) \text{ for all } i \in N, \text{ and } \sigma'_i \in \Sigma_i, \tag{4}$$

and it is a subgame-perfect equilibrium (SPE) if it induces a Nash equilibrium in every subgame, i.e.,

$$U_i(\sigma|h) \geq U_i(\sigma'_i, \sigma_{-i}|h) \text{ for all } i \in N, h \in H, \text{ and } \sigma'_i \in \Sigma_i, \tag{5}$$

where $\sigma|h$ is the restriction of the strategy profile after history $h \in H$.

3.1. Characterization of Equilibria in Repeated Games

The theory of subgame-perfect equilibria in infinitely repeated discounted games with pure strategies has been developed by [1,2,38] (see also [4,10]). The more general model with mixed strategies is analyzed, e.g., in Section 7 in [3].

Let V be the set of subgame-perfect equilibrium payoffs of a repeated game. For any compact set $W \in \mathbb{R}^n$, the punishment payoff of player i in the set W is denoted by

$$p_i(W) = \min\{w_i, w \in W\}.$$

The punishment payoff of player i in the repeated game is the smallest equilibrium payoff, i.e., $p_i(V)$. Let us consider an augmented game where the payoff of each action profile $a \in A$ is given by

$$\tilde{u}_\delta(a) \doteq (I - T)u(a) + Tx(a),$$

where T is a diagonal matrix with $\delta_1, \dots, \delta_n$ on the diagonal. Note that the continuation payoffs $x(a)$ are included in the stage-game payoffs. Let $M(x)$, $x \in \mathbb{R}^{|A| \times n}$, be the set of all equilibrium payoffs in this augmented game. Now, we are ready to state the characterization for the subgame-perfect equilibrium payoffs (see Section 7.3 in [3] for the proof).

Theorem 1. *The set V is the largest bounded fixed point of the mapping B :*

$$W = B(W), \tag{6}$$

where

$$B(W) = \bigcup_{x(a) \in W} M(x),$$

and the stage game's payoffs are given by $(I - T)u(a) + Tx(a)$, $a \in A$.

In contrast to the pure-strategy equilibria, the payoff set V is always non-empty, since every stage game has at least one Nash equilibrium. Note that V is a compact set.

The complexity of computing all mixed-strategy equilibria is much higher compared to the pure-strategy equilibria. Each iteration of B goes through all permutations of all the possible continuations payoffs after each action profile over all action profiles in A .

We denote the payoff set and the mapping B with a discount factor δ by $V(\delta)$ and B^δ . A set W is called self-generating if $W \subseteq B^\delta(W)$. The following result follows directly from Theorem 1 (see also Section 7.3 in [3] for generalizations of these results).

Proposition 1. *If a bounded set W is self-generating then $B^\delta(W) \subseteq V(\delta)$.*

The following result shows that the payoff set is monotone in the discount factor as long as it is convex. For the next two results, we assume that the players have a common discount factor, which is denoted by scalar $\delta = \delta_1 = \dots = \delta_n$.

Theorem 2. *If $V(\delta)$ is convex, then $V(\delta) \subseteq V(\delta')$ for scalar $\delta' \geq \delta$.*

It is also possible to show that convex self-generating sets are monotone. The proof is similar to Theorem 2.

Proposition 2. *If a self-generating set $W \subseteq V(\delta)$ is convex, then $W \subseteq V(\delta')$ for scalar $\delta' \geq \delta$.*

3.2. Self-Supporting Sets and Monotonicity

One problem in finding subgame-perfect equilibria in repeated games is the task of finding the strategy profiles that yield exactly the continuation payoffs such that the players are indifferent between at least two of their pure actions. It turns out that in many games it is possible to find such strategy profiles by selecting the continuation payoffs in such a way that the resulting augmented game corresponds to Game 1 (so with a two-dimensional set of equilibrium payoffs). This means that the play in the repeated game on that round is strategically similar to the play in the one-shot game of Game 1. In the next section, we will show how this construction takes place. It makes use of so-called self-supporting sets that are defined below.

Recall that, for $x \in \mathbb{R}^{|A| \times n}$, the set $M(x)$ consists of all equilibrium payoffs in the augmented game where for each $a \in A$ the continuation payoffs $x(a)$ are included in the stage-game payoffs.

Definition 1. *The set S is a self-supporting set if $S \subseteq M(x)$ for some $x \in \mathbb{R}^{|A| \times n}$ and if for each payoff $s \in S$, (one of) the corresponding equilibrium profile(s) $q(s)$, has the following properties:*

1. $a \in \text{Car}(q(s)) \Rightarrow x(a) \in S$, and
2. if player i plays an action \tilde{a}_i outside $\text{Car}(q_i(s))$ (an observable deviation), while $a_{-i} \in \text{Car}(q_{-i}(s))$, then $x_i(\tilde{a}_i, a_{-i})$ is player i 's punishment payoff $p_i(V)$.
3. if at least two players make an observable deviation, then the continuation payoff is a predetermined equilibrium payoff.

Furthermore, we say that S is a strongly self-supporting set if it does not rely on the punishment payoffs or if the punishment payoffs are included in the set, i.e., $x(a) \in S$ for all $a \in A$. Note that each Nash equilibrium payoff of the stage game by itself forms a strongly self-supporting set. Thus, the self-supporting sets exist in all repeated games but they may not be large or high dimensional.

The difference to the self-generating sets is that the continuation payoffs $x(a)$ on the equilibrium path are generated by a single augmented game (the set $M(x)$) for the self-supporting sets. It also means that the players face the same augmented game on each round, unless there is an observable deviation. Thus, the strategies are simple but not stationary since the players may use different mixed actions after each realized pure-action profile. The strategies on the equilibrium path can be presented with a finite number of states, but they are not of one-period memory [29] (i.e., it is not enough to have the pure-action profiles as states) since some pure-action profile may trigger a punishment in some state but may be acceptable in another state. It should also be noted that the continuation payoffs need not be extreme payoffs, as is the case with correlated strategies and the bang-bang result [3]. The following example [24] shows that self-generating sets are more complicated as they involve multiple stage games with different continuation payoffs.

Example 1. *We examine Prisoner's Dilemma in Figure 2(a) with $\delta = 0.25$, and show that the set consisting of $(1, 1)$ and two line segments $3.5 \times [1.75, 3.5]$ and $[1.75, 3.5] \times 3.5$ is a self-generating set but not self-supporting.*

The line segment $3.5 \times [1.75, 3.5]$ can be generated with Game L1 in Figure 2(b) using continuation payoffs $(3.5, 3.5)$, $(1, 1)$, $(2, 3.5)$ and $(1, 1)$. Game L1 can be found by continuing action profile a by path a^∞ (a corresponds to $(3.5, 3.5)$, b to $(0, 4)$, c to $(4, 0)$ and d to $(1, 1)$), b and d by path d^∞ , and c by a suitable mixed strategy from Game L2 (row player chooses top and column player uses mixed strategy $(3/7, 4/7)$). For example, $(1 - \delta)(4, 0) + \delta(2, 3.5) = (3.5, 0.875)$. All the continuation payoffs belong to the self-generating set itself. The Nash equilibria of Game L1 are the line $3.5 \times [1.75, 3.5]$ and the point $(1, 1)$. Note that the payoff $(2, 3.5)$ is the only one that is outside this set. Game L2 in Figure 2(c) can be found similarly and note that $(2, 3.5)$ is an equilibrium payoff in game L2. Thus, these two lines support each other and they are (together with $(1, 1)$) a self-generating set. This set is not self-supporting as it requires two stage games and two sets of continuation payoffs, and thus this construction is more difficult to find.

The following result shows that certain self-supporting sets are robust in the discount factor if the punishment strategies do not get weaker. This means that the payoffs generated by the self-supporting sets only fill the payoff set more when the players become more patient.

Prisoner's Dilemma	Game L1	Game L2												
<table border="1" style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 5px;">3.5,3.5</td> <td style="padding: 5px;">0,4</td> </tr> <tr> <td style="padding: 5px;">4,0</td> <td style="padding: 5px;">1,1</td> </tr> </table>	3.5,3.5	0,4	4,0	1,1	<table border="1" style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 5px;">3.5,3.5</td> <td style="padding: 5px;">0.25,3.25</td> </tr> <tr> <td style="padding: 5px;">3.5,0.875</td> <td style="padding: 5px;">1,1</td> </tr> </table>	3.5,3.5	0.25,3.25	3.5,0.875	1,1	<table border="1" style="border-collapse: collapse; width: 100%;"> <tr> <td style="padding: 5px;">3.5,3.5</td> <td style="padding: 5px;">0.875,3.5</td> </tr> <tr> <td style="padding: 5px;">3.25,0.25</td> <td style="padding: 5px;">1,1</td> </tr> </table>	3.5,3.5	0.875,3.5	3.25,0.25	1,1
3.5,3.5	0,4													
4,0	1,1													
3.5,3.5	0.25,3.25													
3.5,0.875	1,1													
3.5,3.5	0.875,3.5													
3.25,0.25	1,1													
(a)	(b)	(c)												

Figure 2. Prisoner's dilemma and the stage games used in constructing a self-generating set.

Theorem 3. Let $\delta \in [0, 1)$ and let T be the $n \times n$ diagonal matrix with scalar δ on the diagonal. Let S be a self-supporting set for δ and suppose that the following conditions hold:

- (i) S is a convex set,
- (ii) For all $s \in S$ with corresponding action profile $q(s)$ we have:
 $\tilde{u}_\delta(a) = (I - T)u(a) + Tx(a) \in S$ for all $a \in \text{Car}(q(s))$, and
- (iii) $p_i(V(\delta))$ is not increasing in δ for all $i \in N$.

Also, let $\delta' \geq \delta$ and let T' be the corresponding diagonal matrix. Then there exists a self-supporting set S' for δ' such that $S \subseteq S'$.

Proof. Let S be a self-supporting set for δ , $s \in S$ and $q(s)$ be an equilibrium profile satisfying the three mentioned properties. We show that the same augmented game payoffs can be obtained for discount factor $\delta' \geq \delta$, and this does not introduce any profitable unilateral deviations to the players.

First, we show that there exists a continuation payoff $x'(a) \in S$ for all $a \in \text{Car}(q)$ such that $\tilde{u}_{\delta'}(a) = \tilde{u}_\delta(a)$, or

$$(I - T')u(a) + T'x'(a) = (I - T)u(a) + Tx(a).$$

From this condition, we get

$$T'x'(a) = Tx(a) + (T' - T)u(a)$$

and thus $x'(a)$ is between $x(a)$ and $u(a)$. Substituting $u(a)$ into this equation yields:

$$(I - T)T'x'(a) = (T' - T)\tilde{u}(a) + (I - T')Tx(a).$$

This means that $x'(a)$ is a convex combination of $\tilde{u}(a)$ and $x(a)$. However, since S is convex and $\tilde{u}(a) \in S$ and $x(a) \in S$, this implies that $x'(a) \in S$. Moreover, the continuation payoffs are SPE payoffs with δ' ; they are generated by the set itself.

It remains to show that there are no profitable deviations that are observable. The incentive compatibility conditions hold for δ :

$$(1 - \delta_i)u(q) + \delta_i w_i \geq (1 - \delta_i)d_i(q) + \delta_i p_i(V(\delta)), \tag{7}$$

for all $i \in N$. We show that the same profile q and the continuation payoffs w satisfy these conditions for δ' . Let $\delta'_i = \delta_i + \epsilon_i$, where $\epsilon_i \geq 0$. We have:

$$\begin{aligned} (1 - \delta'_i)u_i(q) + \delta'_i w_i &= (1 - \delta_i)u_i(q) + \delta_i w_i - \epsilon(u_i(q) - w_i) \\ &\geq (1 - \delta_i)d_i(q) + \delta_i p_i(V(\delta)) - \epsilon(u_i(q) - w_i) \\ &= (1 - \delta_i)d_i(q) - \epsilon u_i(q) + \delta_i p_i(V(\delta)) + \epsilon w_i \\ &\geq (1 - \delta_i)d_i(q) - \epsilon d_i(q) + \delta_i p_i(V(\delta)) + \epsilon p_i(V(\delta)) \\ &\geq (1 - \delta'_i)d_i(q) + \delta'_i p_i(V(\delta')) \end{aligned}$$

Here, we used (7) in the first inequality. In the second inequality, we used that $d_i(q) \geq u_i(q)$ (otherwise there would be no reason to deviate) and that $w_i \geq p_i(V(\delta))$ (which follows from $x(a) \geq p_i(V(\delta))$ for all $a \in Car(q)$ by subgame-perfection). Finally, the last inequality follows from $p_i(V(\delta')) \leq p_i(V(\delta))$. Hence, the incentive compatibility constraints are monotone in the discount factor as long as the punishment payoffs do not increase, as observed with pure strategies in Lemma 1 in [10]. \square

In general, the set of equilibrium payoffs needs not be monotone in the discount factor [25,32,39]. However, if the set of continuation payoffs is convex, e.g., if correlated strategies are available, then the payoff set is monotone in the discount factor [2]. Moreover, if the punishment payoffs are included in the set S , i.e., the set is strongly self-supporting, then the self-supporting set is monotone if it is convex.

Proposition 3. *If S is a convex strongly self-supporting set for scalar δ , then there is a convex strongly self-supporting set S' for $\delta' \geq \delta$ such that $S \subseteq S'$.*

We note that Theorem 3 and Proposition 3 could be extended to vector-valued δ if S is additionally assumed to be a hyper-rectangle. This means that the discount factors can be unequal and $\delta' \geq \delta$ means that $\delta'_i \geq \delta_i$ for all $i = 1, \dots, n$. In two-player games, the convex self-supporting sets are always rectangles, which implies that the above results hold for vector-valued δ in these games. The following result gives a sufficient condition for the existence of rectangular self-supporting sets in two-player games.

Proposition 4. *If in a bimatrix game (A, B) player 1 has two actions i_1 and i_2 such that $\underline{b} := \max_j b_{i_1 j} < \min_j b_{i_2 j} =: \bar{b}$ and, similarly, player 2 has two actions j_1 and j_2 such that $\underline{a} := \max_i a_{i j_1} < \min_i a_{i j_2} =: \bar{a}$, then there exists $\delta^* < 1$ such that for all $\delta \geq \delta^*$ there exists a rectangular self-supporting set $[\underline{a}', \bar{a}] \times [\underline{b}', \bar{b}]$, where $\underline{a}' = \max(\underline{a}, p_1(V(\delta)))$ and $\underline{b}' = \max(\underline{b}, p_2(V(\delta)))$.*

Proof. We can create the following 2×2 augmented subgame using actions i_1, i_2, j_1 and j_2 .

\bar{a}, \bar{b}	\underline{a}', \bar{b}
\bar{a}, \underline{b}'	$\underline{a}', \underline{b}'$

The Nash equilibria of the subgame is the rectangle $[\underline{a}', \bar{a}] \times [\underline{b}', \bar{b}]$. The continuation payoffs for the actions on the j_2 column (or i_2 row) are all lower than \bar{a} (\bar{b}) since the payoffs on the column (row) are all higher than \bar{a} (\bar{b}). Similarly, the continuation payoffs for the actions on the j_1 column (or i_1 row) are all higher than \underline{a}' (\underline{b}') since the payoffs are all lower than \underline{a}' (\underline{b}'). Thus, the continuation payoffs belong to the set itself when the discount factor is high enough. Moreover, the players do not have any incentive to deviate outside these two actions when the discount factor is high enough and the deviations are followed by the punishment payoffs that are smaller than or equal to \underline{a}' or \underline{b}' . \square

4. Strategies in the Repeated Prisoner's Dilemma

In this section, we want to demonstrate the use of self-supporting sets in proving results in repeated games. The results are well-known (see, for example, [27,30,31]). Recall that in symmetric prisoner's dilemmas all players have the same discount factor denoted by scalar δ .

PD	
3, 3	0, 4
4, 0	1, 1

Theorem 4 ([30]). *In game PD above, the set of subgame-perfect equilibrium payoffs for $\delta = 1/3$ is given by the rectangle $[1, 3] \times [1, 3]$ and the two line segments $[3, 11/3] \times 1$ and $1 \times [3, 11/3]$.*

Proof. We first show that the rectangle is a strongly self-supporting set and thus part of the payoff set. Then we show how to get the lines as SPE payoffs and that there are no other SPE payoffs. These results have been developed in [30]⁴, but here we want to demonstrate the new methodology with an example.

We can create an augmented game from game PD that corresponds to Game 1 by choosing the expected continuation payoffs in the following way: let us denote the action pairs by a, b, c and d , where a corresponds to the payoff (3,3), b to (0,4), c to (4,0) and d to (1,1). If the continuation payoffs are $x(a) = (3,3)$, $x(b) = (3,1)$, $x(c) = (1,3)$, and $x(d) = (1,1)$, then the payoffs in the augmented game are the same as the payoffs in Game 1. For example, after action profile b the payoff is $\tilde{u}(b) = (1 - \delta)(0,4) + \delta(3,1) = (0, 8/3) + (1, 1/3) = (1, 3)$. Thus, if the action profile b is realized, then the players continue by playing a strategy pair that gives them a continuation payoff of (3, 1). Now, we show what these 'continuation strategy pairs' look like for each action profile⁵.

The action profiles a and d are straightforward. The continuation payoff (3,3) can be obtained by playing action pair a infinitely. We denote this play by a^∞ , which is also called a path. So, if an action pair a is realized then the continuation strategy pair is a^∞ , and any deviation by any player will be detected and is punished (or followed) by the path d^∞ . Similarly, if d is realized, then the continuation payoff (1, 1) will be obtained by playing the path d^∞ . Finally, the continuation payoff (3,1) (or (1,3)) can be obtained by playing the action pair c (or b) in Game 1. Thus, in the repeated game, the play will alternate deterministically between b and c if one of these action pairs is realized in the augmented game at stage 1. We denote this play by $(cb)^\infty$ after b and $(bc)^\infty$ after c . Note that the equilibrium payoffs in Game 1 form the rectangle $[1, 3] \times [1, 3]$ and all the required continuation payoffs are contained in this set as well. So, $[1, 3] \times [1, 3]$ is actually a strongly self-supporting set, since it even contains the punishment payoffs.

Now, we show that the payoffs on the line segment $[3, 11/3] \times 1$ can be obtained by subgame-perfect equilibria (a similar argument holds for the line segment $1 \times [3, 11/3]$). We will show that we can create an augmented game at stage 1 of game PD that corresponds to Game 2. To do this, we let the action pairs a, b and d be followed by the path d^∞ , and c be followed by the path a^∞ . Then, e.g., $\tilde{u}(c) = (1 - \delta)(4,0) + \delta(3,3) = (11/3, 1)$. However, the (3,3) continuation payoff is not in the set $[3, 11/3] \times 1$. This means that the line segment is not a self-supporting set: it requires a continuation payoff that is outside the set.

It remains to show that there are no other subgame-perfect equilibrium payoffs. Assume on the contrary that there is a subgame-perfect equilibrium payoff $(3 + w_1, 1 + w_2)$, where $w_1, w_2 > 0$. It is easy to check that this payoff can only be obtained by playing first the action profile c . To generate the payoff, the continuation payoff (x_1, x_2) must satisfy $\tilde{u}(c) = (1 - \delta)(4,0) + \delta(x_1, x_2) = (3 + w_1, 1 + w_2)$. From this, we get $x_1 = 1 + 3w_1$ and $x_2 = 3 + 3w_2$. This continuation payoff itself can only be

⁴ We thank Tadashi Sekiguchi for pointing this out.

⁵ Note that in general the continuation play involves mixed strategies even though pure strategies are enough in this example. Figure 3b shows an example where mixed actions are used after action profiles b and c .

obtained by playing action profile b first. The required continuation payoff after b turns out to be $(3 + 9w_1, 1 + 9w_2) > (3 + w_1, 1 + w_2)$. We can conclude that this process requires higher and higher continuation payoffs, which will eventually be outside the convex hull of the stage-game payoffs, and thus they cannot be generated by the repeated game. \square

Figure 3a illustrates a strongly self-supporting set in game PD with $\delta = 2/5$, and b the corresponding strategies with a finite number of states. Figure 3a shows that the continuation payoff $x(c)$ belongs to the shaded self-supporting set, and $v = (1.5, 1.5)$ is an arbitrary point in the set. It holds that $\tilde{u}(c) = (1 - \delta)u(c) + \delta x(c)$, which geometrically means that $\tilde{u}(c)$ is $\delta = 0.4$ fraction on the line from $u(c)$ to $x(c)$. When $\delta = 1/3$, $x(c)$ moves to the corner point $(1, 3)$, which is the smallest δ value when the self-supporting set exists. In Figure 3b, we can see the equilibrium strategy pair that produces payoff v . Each circle corresponds to a state with the expected payoffs and the action pair that is played in that state. The states that the arrows point to correspond to the action pairs that should be played next. The probabilities of top and left actions are given by p and q for the row and column players, respectively. The star $*$ denotes that the corresponding transition could be to any equilibrium strategy, since it corresponds to the case where both players deviate from the equilibrium strategy simultaneously. Note that the continuation play involves mixed strategies in general, unless the continuation payoff happens to be one of the pure payoffs.

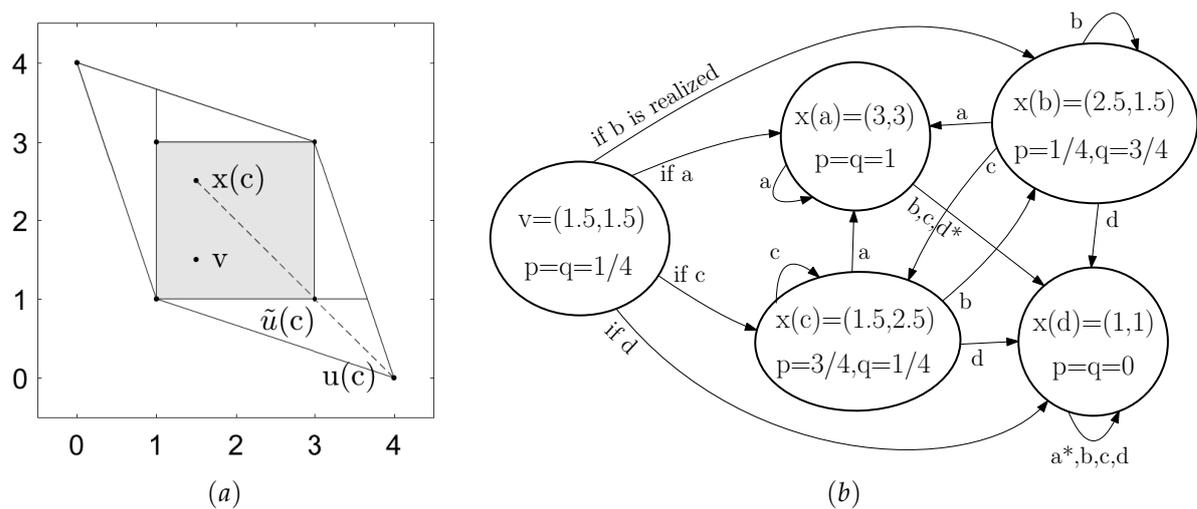


Figure 3. Geometric illustration of (a) a self-supporting set and (b) the strategies in game PD. The star $*$ denotes that the corresponding transition could be to any equilibrium strategy.

Theorem 5. In game PD, for $\delta \geq 1/3$, the rectangle $[1, 3] \times [1, 3]$ is a subset of the set of subgame-perfect equilibrium payoffs.

Proof. Using the fact that $[1, 3] \times [1, 3]$ is a convex strongly self-supporting set, this follows directly from combining Proposition 3 or Theorem 3 with the proof of Theorem 4. \square

$$\begin{array}{c}
 \text{PD2} \\
 \begin{array}{|c|c|}
 \hline
 a, a & b, c \\
 \hline
 c, b & d, d \\
 \hline
 \end{array} \\
 \text{with } c > a > d > b
 \end{array}$$

Theorem 6. In game PD2, a symmetric prisoner’s dilemma, the rectangle $[\vartheta, a] \times [\vartheta, a]$ is a subset of the set of subgame-perfect equilibrium payoffs for

$$\delta \geq \max \left[\frac{c - a}{c - \vartheta}, \frac{\vartheta - b}{a - b} \right].$$

Proof. We want to find the smallest discount factor for which the set $[\vartheta, a] \times [\vartheta, a]$ is strongly self-supporting. This means that the continuation payoff $x(c)$ after action pair c must satisfy $\tilde{u}(c) = (1 - \delta)(c, b) + \delta(x_1(c), x_2(c)) = (a, \vartheta)$. This gives

$$(x_1(c), x_2(c)) = \left(\frac{a - (1 - \delta)c}{\delta}, \frac{\vartheta - (1 - \delta)b}{\delta} \right),$$

where, in order to make the set self-supporting, we require that $x_1(c) \geq \vartheta$ and $x_2(c) \leq a$, leading to $\delta \geq \max [(c - a)/(c - \vartheta), (\vartheta - b)/(a - b)]$. □

5. Example of a Duopoly Game

Let us examine the duopoly model of Abreu [38], where two firms decide their production quantities. The firms have three output levels: low (L), medium (M) and high (H) (see Figure 4). The stage game’s Nash equilibrium is (M,M), which gives the players the payoff (7, 7). The players may obtain many other equilibrium payoffs in the repeated game, including the payoff (10, 10) which is obtained by playing (L,L) repeatedly. This outcome is supported by the punishment⁶ of playing H, which makes the other player obtain, at most, a payoff of 0. In this paper, we show that large sets of payoffs can be obtained using self-supporting sets.

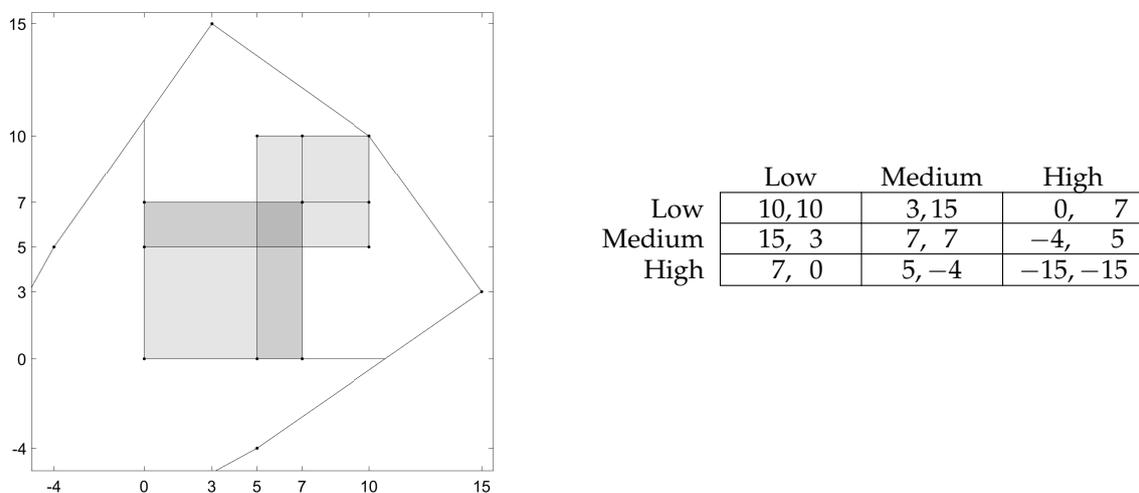


Figure 4. Four self-supporting sets in a duopoly game with three production quantities.

Figure 4 shows four self-supporting sets that exist when the players are patient enough: the squares $[0, 7] \times [0, 7]$ and $[7, 10] \times [7, 10]$, and the rectangles $[0, 10] \times [5, 7]$ and $[5, 7] \times [0, 10]$. These are equilibrium payoffs and they cover almost half of the feasible and individually rational payoffs. The equilibrium payoffs depend on the discount factors and it is difficult to find all of them, however, these self-supporting sets capture majority of them.

⁶ The optimal pure punishment strategies in this game depend on the discount factors [26].

The self-supporting sets in games with more actions can be found by splitting the game into smaller 2×2 games. Randomizing between more than two actions does not bring any new payoffs for the self-supporting sets, and it is more difficult to find the suitable continuation payoffs in the set such that the players are indifferent between all the actions. In this game, the square $[0, 7] \times [0, 7]$ can be found in the No Conflict game in Figure 5a. The square $[7, 10] \times [7, 10]$ can be found in Prisoner's Dilemma game in Figure 5b. The rectangle $[0, 10] \times [5, 7]$ can be found in Game D1 in Figure 5c, and the other rectangle symmetrically by switching the role of the players.

		No Conflict		Prisoner's Dilemma		Game D1			
		L	H	L	M	L	H		
L		10,10	0,7	L	10,10	3,15	L	10,10	0,7
H		7,0	-15,-15	M	15,3	7,7	M	15,3	-4,5
		(a)		(b)		(c)			

Figure 5. Subgames used in constructing the self-supporting sets in the duopoly game.

Note that it is easy to find the self-supporting sets in 2×2 games by Proposition 4. If the minimum payoff against one action is higher than the maximum payoff against the other action and this holds for both players, then the self-supporting set contains all the payoffs between the maximum and the minimum payoffs. In Game D1, the minimum payoff of player 1 against L is 10 and the maximum against H is 0, and the minimum payoff of player 2 against L is 7 and the maximum payoff against M is 5. Thus, the self-supporting set is $[0, 10] \times [5, 7]$. It is easy to check geometrically that the continuation payoffs belong to the set when the discount factors are high enough.

6. Conclusions

In this paper, we construct subgame-perfect mixed-strategy profiles in repeated games. Basically, the players consider on each round an augmented game that takes into account the continuation payoffs that are defined by the players' follow-up strategies. The continuation play and payoffs can be different after each realized pure-action profile but all the continuation payoffs need to be equilibrium payoffs themselves. Moreover, the equilibrium payoffs can be found by computing the Nash equilibria in all these augmented games.

How to compute the mixed-strategy equilibria efficiently is still an open problem. It is more difficult to compute all the mixed-strategy equilibria compared to the pure (or correlated) strategies. However, we have found that the concept of self-supporting sets may facilitate our task considerably, since they may give us large and high-dimensional sets of equilibrium payoffs with relatively simple accompanying strategy profiles. It is left to future research to determine how common large self-supporting sets are in multiplayer games and under which conditions they do exist. Although, it may be very hard to find all the equilibrium payoffs and corresponding strategy profiles, it may be possible to compute some payoffs or good approximations fast. We think that our concept of self-supporting sets may open up ideas for new simple strategies that produce large sets of payoffs in repeated and stochastic games.

Acknowledgments: Kimmo Berg acknowledges funding from Emil Aaltosen Säätio through Post doc -pooli.

Author Contributions: Kimmo Berg and Gijs Schoenmakers developed the main idea and wrote the paper together. Kimmo Berg developed the numerical examples and illustrations.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Abreu, D.; Pearce, D.; Stacchetti, E. Optimal cartel equilibria with imperfect monitoring. *J. Econ. Theory* **1986**, *39*, 251–269.
2. Abreu, D.; Pearce, D.; Stacchetti, E. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* **1990**, *58*, 1041–1063.
3. Mailath, G.J.; Samuelson, L. *Repeated Games and Reputations: Long-Run Relationships*; Oxford University Press: Oxford, UK, 2006.
4. Cronshaw, M.B.; Luenberger, D.G. Strongly symmetric subgame perfect equilibria in infinitely repeated games with perfect monitoring. *Games Econ. Behav.* **1994**, *6*, 220–237.
5. Cronshaw, M.B. Algorithms for finding repeated game equilibria. *Comput. Econ.* **1997**, *10*, 139–168.
6. Judd, K.; Yeltekin, Ş.; Conklin, J. Computing supergame equilibria. *Econometrica* **2003**, *71*, 1239–1254.
7. Burkov, A.; Chaib-draa, B. An approximate subgame-perfect equilibrium computation technique for repeated games. In Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, Atlanta, GA, USA, 11–15 July 2010; pp. 729–736.
8. Salcedo, B.; Sultanum, B. *Computation of Subgame-Perfect Equilibria of Repeated Games with Perfect Monitoring and Public Randomization*; Working Paper; The Pennsylvania State University: State College, PA, USA, 2012.
9. Abreu, D.; Sannikov, Y. An algorithm for two player repeated games with perfect monitoring. *Theor. Econ.* **2014**, *9*, 313–338.
10. Berg, K.; Kitti, M. *Equilibrium Paths in Discounted Supergames*; Working Paper; Aalto University: Espoo, Finland, 2012. Available online: <http://sal.aalto.fi/publications/pdf-files/mber09b.pdf> (accessed on 24 October 2017).
11. Berg, K.; Kitti, M. Computing equilibria in discounted 2×2 supergames. *Comput. Econ.* **2013**, *41*, 71–78.
12. Berg, K.; Kitti, M. Fractal geometry of equilibrium payoffs in discounted supergames. *Fractals* **2014**, *22*.
13. Fudenberg, D.; Maskin, E. The folk theorem in repeated games with discounting and incomplete information. *Econometrica* **1986**, *54*, 533–554.
14. Fudenberg, D.; Maskin, E. On the dispensability of public randomization in discounted repeated games. *J. Econ. Theory* **1991**, *53*, 428–438.
15. Gossner, O. The Folk Theorem for Finitely Repeated Games with Mixed Strategies. *Int. J. Game Theory* **1995**, *24*, 95–107.
16. Busch, L.-A.; Wen, Q. Negotiation games with unobservable mixed disagreement actions. *J. Math. Econ.* **2001**, *35*, 563–579.
17. Fudenberg, D.; Levine, D.; Maskin, E. The folk theorem with imperfect public information. *Econometrica* **1994**, *62*, 997–1039.
18. Sugaya, T. Characterizing the limit set of PPE payoffs with unequal discounting. *Theor. Econ.* **2015**, *10*, 691–717.
19. Chassang, S.; Takahashi, S. Robustness to incomplete information in repeated games. *Theor. Econ.* **2011**, *6*, 49–93.
20. Peski, M. Repeated games with incomplete information and discounting. *Theor. Econ.* **2014**, *9*, 651–694.
21. Dutta, P.K. A Folk theorem for stochastic games. *J. Econ. Theory* **1995**, *66*, 1–32.
22. Hörner, J.; Sugaya, T.; Takahashi, S.; Vieille, N. Recursive methods in discounted stochastic games: An algorithm for $\delta \rightarrow 1$ and a folk theorem. *Econometrica* **2011**, *79*, 1277–1318.
23. Berg, K. Elementary subpaths in discounted stochastic games. *Dyn. Games Appl.* **2016**, *6*, 304–323, doi:10.1007/s13235-015-0151-5.
24. Berg, K. *Set-Valued Games and Mixed-Strategy Equilibria in Discounted Supergames*; Working Paper; Aalto University: Espoo, Finland, 2017.
25. Berg, K. Extremal pure strategies and monotonicity in repeated games. *Comput. Econ.* **2017**, *49*, 387–404, doi:10.1007/s10614-016-9565-4.
26. Berg, K.; Kärki, M. *An Algorithm for Computing the Minimum Pure-Strategy Payoffs in Repeated Games*; Working Paper; Aalto University: Espoo, Finland, 2016. Available online: <http://sal.aalto.fi/publications/pdf-files/mber14c.pdf> (accessed on 24 October 2017).
27. Ely, J.C.; Välimäki, J. A Robust Folk Theorem for the Prisoner’s Dilemma. *J. Econ. Theory* **2002**, *102*, 84–105.
28. Ely, J.C.; Hörner, H.; Olszewski, W. Belief-free equilibria in repeated games. *Econometrica* **2005**, *73*, 377–415.
29. Doraszelski, U.; Escobar, J.F. Restricted feedback in long term relationships. *J. Econ. Theory* **2012**, *147*, 142–161.

30. Sorin, S. On repeated games with complete information. *Math. Oper. Res.* **1986**, *11*, 147–160.
31. Stahl, D.O. The graph of prisoner's dilemma supergame payoffs as a function of the discount factor. *Games Econ. Behav.* **1991**, *3*, 368–384.
32. Mailath, G.J.; Obara, I.; Sekiguchi, T. The maximum efficient equilibrium payoff in the repeated prisoners' dilemma. *Games Econ. Behav.* **2002**, *40*, 99–122.
33. Berg, K.; Kärki, M. *How Patient the Players Need to Be to Obtain All the Relevant Payoffs in Discounted Supergames?* Working Paper; Aalto University: Espoo, Finland, 2014. Available online: <http://sal.aalto.fi/publications/pdf-files/mber14b.pdf> (accessed on 24 October 2017).
34. Bhaskar, V.; Mailath G.J.; Morris, S. Purification in the infinitely-repeated prisoners' dilemma. *Rev. Econ. Dyn.* **2008**, *11*, 515–528.
35. Bhaskar, V.; Mailath G.J.; Morris, S. A foundation for Markov equilibria in sequential games with finite social memory. *Rev. Econ. Stud.* **2013**, *80*, 925–948.
36. Robinson, D.; Goforth, D. *The Topology of the 2×2 Games: A New Periodic Table*; Routledge: Abingdon, UK, 2005.
37. Borm, P. A classification of 2×2 bimatrix games. *Cah. Cent. d'Etudes Rech. Oper.* **1987**, *29*, 69–84.
38. Abreu, D. On the theory of infinitely repeated games with discounting. *Econometrica* **1988**, *56*, 383–396.
39. Yamamoto, Y. The use of public randomization in discounted repeated games. *Int. J. Game Theory* **2010**, *39*, 431–443.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).