Repeated games with voluntary information purchase

János Flesch, Andrés Perea *

Maastricht University, Department of Quantitative Economics, PO Box 616, 6200 MD Maastricht, The Netherlands

Abstract

We consider discounted repeated games in which players can voluntarily purchase information about the opponents' actions at past stages. Information about a stage can be bought at a fixed but arbitrary cost. Opponents cannot observe the information purchase by a player. For our main result, we make the usual assumption that the dimension of the set *FIR* of feasible and individually rational payoff vectors is equal to the number of players. We show that, if there are at least three players and each player has at least four actions, then every payoff vector in the interior of the set *FIR* can be achieved by a Nash equilibrium of the discounted repeated game if the discount factor is sufficiently close to 1. Therefore, nearly efficient payoffs can be achieved even if the cost of monitoring is high. We show that the same result holds if there are at least four players and at least three actions for each player. Finally, we indicate how the construction can be extended to sequential equilibrium.
© 2008 Elsevier Inc. All rights reserved.

JEL classification: C73

Keywords: Repeated games; Folk Theorem; Imperfect monitoring; Costly monitoring

1. Introduction

The Folk Theorem in repeated games states that, if players can perfectly and costlessly observe the opponents' past actions, and evaluate payoff streams by means of the limiting average criterion, then every feasible and individually rational payoff vector of the stage game can be obtained by a Nash equilibrium in the repeated game (see Aumann and Shapley, 1994). Fudenberg and Maskin (1986) extended this result to the discounted case under the condition that the dimension of the set *FIR* of feasible and individually rational payoff vectors is equal to the number of players (for later reference, we call this the *full dimensionality condition*[1]). The proofs heavily rely on the assumption that information about the opponents' behavior can be obtained costlessly. Recently, there has been a growing interest in models of discounted repeated games in which such information is no longer costless. See, for instance, Ben-Porath and Kahneman (2003), Miyagawa et al. (2004) and Kandori and Obara (2004). These papers all study models in which players can voluntarily monitor the actions of a group of opponents at the current stage. They differ, however, on many

---

* Corresponding author at: Maastricht University, Department of Quantitative Economics, PO Box 616, 6200 MD Maastricht, The Netherlands.

E-mail address: a.perea@ke.unimaas.nl (A. Perea).

[1] This condition, which has been widely used, merely rules out degenerate cases.

assumptions regarding the number of players, the quality of monitoring, the information that players receive if they do not monitor, the availability of costless public messages, and other aspects.

Our paper comes closest to Ben-Porath and Kahneman (2003) in terms of the model, the results and the construction used to prove the result, but also bears some resemblance with Miyagawa et al. (2004). For this reason, we will compare our construction with Ben-Porath and Kahneman and Miyagawa et al. throughout this paper.

In our paper, we analyze infinitely repeated games with discounting in which purchasing information about opponents' actions is voluntary but costly. More precisely, at the beginning of every stage a player can choose a subset of past stages at which he wants to buy information. For each stage he wishes to buy, he must pay a fixed amount $c > 0$, and will observe the actions of every opponent at that stage. The net discounted payoff for a player is given by his usual expected discounted payoff, minus the expected discounted cost of buying information. Our main result is to show that, if full dimensionality holds, the number of players is at least three, and each of them has at least four actions, then every payoff vector in the interior of *FIR* can be supported (in terms of net expected discounted payoffs) by a Nash equilibrium, given that the discount factor is sufficiently close to 1. Besides, we show that the same result holds if there are at least four players and at least three actions for each player. We also indicate how the result can be extended to sequential equilibrium.

Our model differs from Ben-Porath and Kahneman (2003) and Miyagawa et al. (2004) on the following aspects: (1) In our setting, a player can also purchase information about previous stages, whereas in Ben-Porath and Kahneman and Miyagawa et al. a player can only buy information about the present stage; (2) If a player decides to buy information in our model, he observes the actions of all opponents, whereas in Ben-Porath and Kahneman and Miyagawa et al. a player can choose the group of opponents which he would like to monitor; (3) In our model and in Miyagawa et al. public announcements are not available, whereas Ben-Porath and Kahneman assumes costless public announcements; (4) In our model and in Ben-Porath and Kahneman, a player does not receive any information about the opponents' actions unless he buys information. In Miyagawa et al., even if a player does not buy information he receives a costless stochastic signal depending on the action profile chosen at the current stage; (5) In our construction we do not use any public randomization device, in contrast with Miyagawa et al. In Ben-Porath and Kahneman, such a device is only used to simplify the construction. In conclusion, we would like to emphasize that we assume no external devices (for public announcements or randomization), in contrast with other models. This makes the communication and coordination of the players more troublesome in our context.

All the models above can be applied to several economic environments of interest. Possible applications include repeated partnership models (see, for instance, Radner, 1986 and Radner et al., 1986) or models of collusion between firms (for instance Green and Porter, 1984).

We shall now provide a motivation for the construction of our Nash equilibria in the repeated game. Take an arbitrary payoff vector in the interior of *FIR*. If it cannot be supported by a convex combination of Nash equilibria of the stage game, then at certain stages players must play action profiles where at least one of them would have an incentive to deviate in terms of stage payoffs. It is therefore essential that this player will be monitored with positive probability at such stages. However, in order to give an opponent an incentive to purchase information about these stages, he must have some uncertainty about the players' actions there. This may be achieved by the following delicate construction: Choose a large number n, and a sequence of 2^n pure action profiles such that the corresponding discounted payoff is close to the desired payoff vector. During a block U of 2^n stages, players are to follow these action profiles. However, at the beginning of U each player chooses a stage at which he will "switch" from the prescribed action, according to a cleverly chosen probability distribution putting a positive probability on each of these 2^n stages. We will come back to this probability distribution towards the end of this introduction.

Since every stage in U is chosen with positive probability as a possible switching stage, we must guarantee that each of these switching stage choices leads to the same expected payoff for every player. As to achieve this we build in a compensation mechanism by making the future target payoff dependent on the switching stages chosen and the realized actions at these switching stages. For this, we let the players report their switching stage after block U, thereby enabling the opponents to discover all switching stages and the corresponding switching actions. This means that all realized actions in U will eventually be known to all players. Recall, however, that public announcements are not available in our setting, which makes the exchange of information more troublesome. Players must use regular actions to report their switching stages, which can be discovered by the opponents by purchasing information. We construct after block U a block I with n stages in which players report their switching stage in binary code. For each

player, two actions are used to code digit "1," and two other actions are used to code digit "0" (recall that we assume at least four actions for every player). This way players are able to randomize while sending information.

In order to give the players an incentive to buy all stages in I, and subsequently to buy the reported switching stages, we construct a block C, with only one stage, after I in which every player must make a report based on the observed actions at the switching stages. By construction, we make sure that if a player decides not to buy information when he is supposed to do so, then he will report incorrectly with a positive probability, which would then be detected by his opponents. Finally, there is a block D, with only one stage, after C in which players can report whether they have detected a deviation from the prescribed strategy profile. If a deviation is reported, all players can discover the deviator by purchasing past stages, and this deviator will be punished for the remainder of the game.

After block D, a new cycle of blocks U, I, C and D will follow, and so on. For the new block U, a new target payoff vector is chosen based on the observed actions in the previous cycle of blocks. In the new block C, players must report the actions they observed in the previous blocks C and D. This is done in order to give the players an incentive to purchase information about the blocks C and D.

The outline of this paper is as follows: In Section 2 we present the model. In Section 3 we present our main result, and give a detailed outline of the proof. In Section 4 we give the formal proof for our main result. In Section 5 we give the construction for sequential equilibrium. Finally, in Section 6 we discuss possible extensions of the main result, and give some concluding remarks.

2. Model

We consider repeated games in which the same stage game is repeated infinitely often, and players evaluate payoff streams by means of the same[2] discount factor. At every stage, a player has the opportunity to purchase information about the past actions chosen by his opponents. If he decides to buy information, he has to pay a fixed price[3] for every past stage at which he wants the opponents' behavior to be revealed. At a given stage, a player may purchase information about as many past stages as he likes.

Formally, the *stage game* is a simultaneous-move game $\Gamma^s = ((A_i)_{i \in I}, (v_i)_{i \in I})$, where $I = \{1, \dots, m\}$ is the finite set of players, A_i is the finite set of pure actions for player i, and $v_i : \times_{j \in I} A_j \to \mathbb{R}$ is player i's payoff function. At every stage $t \in \{1, 2, 3, \dots\}$ of the repeated game, player i first chooses a (possibly empty) collection $K_i^t(1) \subseteq \{1, 2, \dots, t-1\}$ of past stages about which he wants to buy information. Subsequently, he observes the opponents' actions that were chosen at stages in $K_i^t(1)$, and depending on this information he may decide to purchase another collection $K_i^t(2)$ of past stages not already included in $K_i^t(1)$, and so on, until he does not desire any further information.[4] Let $\bar{K}_i^t$ be the collection of all stages about which player i buys information at stage t. For each stage he buys, player i must pay a fixed amount $c > 0$, resulting in a total cost of $c_i^t = c|\bar{K}_i^t|$. After collecting all this information, player i chooses a pure action $a_i^t \in A_i$. If every player does so at stage t, this results in an action profile $a^t = (a_j^t)_{j \in I}$ which yields every player i a payoff $v_i^t = v_i(a^t)$. The infinite sequence $((v_i^1, c_i^1), (v_i^2, c_i^2), \dots)$ of payoffs and costs for player i is evaluated by means of the δ-discounted payoff

$$u_i^\delta = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} (v_i^t - c_i^t),$$

where $\delta \in (0, 1)$ is the common discount factor.

We assume that every player (1) knows the stage game, the price c for information and the common discount factor, (2) observes his own action,[5] but does not observe the opponents' actions at a given stage t unless he buys

---

information about stage t, (3) is unaware of the information purchase by his opponents at a given stage t, even if he buys information about stage t, and (4) has perfect recall.

A *behavior strategy* σ_i for player i specifies, for every stage and every possible collection of past actions[6] that may have been observed until then, (1) an information purchase as described above, and, based on this new information, (2) a probability distribution over pure actions. A collection $\sigma = (\sigma_i)_{i \in I}$ is called a behavior strategy profile, and induces for every player i an expected δ-discounted payoff $u_i^\delta(\sigma)$.

A behavior strategy profile $\sigma = (\sigma_i)_{i \in I}$ is a *Nash equilibrium* for discount factor $\delta \in (0, 1)$ if $u_i^\delta(\sigma) = \max_{\sigma_i'} u_i^\delta(\sigma_i', \sigma_{-i})$. Here, σ_{-i} is a short way to write $(\sigma_j)_{j \neq i}$. Let $E^\delta \subseteq \mathbb{R}^I$ denote the set of δ-discounted payoff vectors that correspond to Nash equilibria for the discount factor δ.

Let $A = \times_{i \in I} A_i$ be the set of pure action profiles in the stage game Γ^s. By $F = \text{conv}\{(v_i(a))_{i \in I} | a \in A\} \subseteq \mathbb{R}^I$ we denote the set of feasible payoffs in the stage game, where "conv" stands for "convex hull." Let

$$d_i = \min_{\alpha_{-i} \in \times_{j \neq i} \Delta(A_j)} \max_{a_i \in A_i} v_i(a_i, \alpha_{-i})$$

be the min-max payoff for player i in the stage game. Here, $\Delta(A_j)$ denotes the set of probability distributions on A_j and $v_i(a_i, \alpha_{-i})$ is the expected payoff for player i induced by a_i and α_{-i}. By

$$FIR = \{(v_i)_{i \in I} \in F \mid v_i \geqslant d_i \text{ for every } i\}$$

we denote the set of feasible and individually rational payoff vectors in the stage game.

3. Main result and detailed outline of proof

3.1. Main result

Take a stage game with a full dimensional *FIR*, and assume that there are at least three players and that each player has at least four actions. Our main result is to show that every payoff vector in the interior of *FIR* is induced by some equilibrium in the δ-discounted repeated game if δ is sufficiently close to 1. In Section 5 we will discuss possible generalizations of this result, including a construction for sequential equilibrium.

Theorem 3.1. *Let Γ^s be a stage game with a full dimensional FIR. Assume that there are at least three players and at least four actions for each player. Then, for every $u \in \text{int}(FIR)$ there is some $\delta(u) \in (0, 1)$ such that $u \in E^\delta$ for every $\delta \in (\delta(u), 1)$.*

Without entering into too much detail here, we would like to clarify the assumptions we make in our theorem. The reason we assume at least four actions for every player is to enable the players to communicate through randomizations over actions. We need at least three players since in our construction it is essential that each player's actions are monitored by at least two opponents as to make deviations detectable (see DM2 in Section 3.3).[7]

Before presenting the formal proof in the following section, we illustrate the main idea by means of an example. Consider the following three-player stage game:

	a_2	b_2			a_2	b_2
a_1	1, 1, 1	0, 0, 0		a_1	0, 0, 0	3, 0, 0
b_1	0, 0, 0	0, 0, 3		b_1	0, 3, 0	2, 2, 2
	a_3				b_3	

Here, player 1 chooses the row, player 2 chooses the column and player 3 chooses the matrix. Assume that every player i has two additional actions $\bar{a}_i, \bar{b}_i$ which are simply duplicates of a_i and b_i (that is, $\bar{a}_i$ induces the same payoffs

---

that in generic games observing your own payoffs would imply full monitoring, in which case the main theorem would hold without any need to buy additional information (see Fudenberg and Maskin, 1986).

[6] One could also assume that behavior strategies take into account *when* information purchase took place. For our purposes it makes no difference.

[7] Also Ben-Porath and Kahneman (2003) treat the case of two players as a special one.

as a_i against any pair of actions by the opponents, and similarly for $\bar{b}_i$). Notice that $d_i = 0$ for all players i, and that $\dim(FIR) = 3$. Hence, the conditions of our theorem are met.

Take a payoff vector u in the interior of FIR. We show how to construct an equilibrium for large enough δ which induces u.

3.2. The strategy profile

The first step is to divide the set of stages into "master-blocks" $M_1, M_2, \ldots$ of length $2^n + n + 2$ (where n is yet to be determined). Master-block M_1 is designed in such a way that the realized payoff-vector in M_1 is close to u. Since in M_1 the strategy profile does not always prescribe a fixed action, one cannot guarantee that the realized payoff-vector is exactly equal to u. Master-block M_2 is identical to M_1. By construction, at the end of M_2 the players will find out the realized payoffs in block M_1. Before the start of M_3, a new target-payoff-vector u_3 has to be determined as to compensate for the "contaminating" payoffs in M_1. In general, before the start of each master-block M_k we determine a target-payoff-vector u_k based on the difference between u_{k-2} and the realized payoffs in M_{k-2}. An important step in our construction will be to make sure that the target-payoff-vector u_k sufficiently stays away from the boundary of FIR. The reason for not allowing u_k to be close to the boundary of FIR is that players must be able to randomize at certain stages in the construction.

Each master-block M_k consists of four sub-blocks, the "utility-block" U_k (of length 2^n), the "information-exchange-block" I_k (of length n), the "control-block" C_k (of length 1), and the "deviation-report-block" D_k (of length 1). Table 1 summarizes the players' behavior in these four sub-blocks.

Utility-block U_k: This block makes sure that the total discounted payoff-vector converges to u, as desired. In order to achieve this, the relative length of U_k within the master-block M_k is chosen close to 1 (by choosing n large enough).

Information purchase: During block U_k no player buys any information yet.

Actions: The prescribed actions here are pure actions in $\{a_i, b_i\}$, and are chosen in such a way that the induced discounted payoff-vector is close to the target-payoff-vector u_k. However, every player i chooses a stage in this utility-block in which he "switches" from the prescribed action to randomizing equally between the two duplicate actions $\bar{a}_i$ and $\bar{b}_i$. This switching-stage is chosen according to the following probability distribution: the last stage in U_k is chosen with probability $2^n/(2^n + c)$, whereas the other stages are chosen with probability $c/(2^n - 1)(2^n + c)$. Hence, the last stage is chosen with a much larger probability than the other stages in U_k.[8]

Table 1
Summary of player i's behavior as long as no player has reported a deviation

	Actions	Information purchase

U_k
 (2^n stages)
	∗ choose switching-stage $t_{ik} \in U_k$ as follows:
	   last stage in U_k with probability $2^n/(2^n + c)$
	   all other stages in U_k with equal probability
	 ∗ at all stages except t_{ik} play prescribed pure action
	 ∗ at stage t_{ik} switch to $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$
		none

I_k (n stages)	report switching-stage t_{ik} in binary code	none
C_k (1 stage)	∗ play $\frac{1}{2}a_i + \frac{1}{2}b_i$ if the sum of the number of a's and the number of $\bar{a}$'s at t_{jk} ($\forall j$), C_{k-1}, D_{k-1} is even ∗ play $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ otherwise	∗ I_k ∗ t_{jk} for all $j \neq i$ ∗ C_{k-1} ∗ D_{k-1}
D_k (1 stage)	∗ if player i detects a deviation then report this by playing $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ ∗ play $\frac{1}{2}a_i + \frac{1}{2}b_i$ otherwise	none

---

[8] In Ben-Porath and Kahneman (2003), a similar block is used to generate a discounted payoff close to the target payoff. A difference is that in Ben-Porath and Kahneman, every player randomly chooses *two* switching stages, according to the *uniform* distribution. Also, in Ben-Porath and Kahneman every player monitors exactly one opponent at a randomly chosen stage. In Ben-Porath and Kahneman, namely, buying information about opponents' actions cannot be delayed. The monitoring assignment in Ben-Porath and Kahneman is quite delicate.

Information-exchange-block I_k: The purpose of I_k is to enable the players to find out the realization of the switches at U_k.

Information purchase: During block I_k no player buys any information yet.

Actions: Each player must report in binary code the stage in U_k at which he switched. Here, randomization $\frac{1}{2}a_i + \frac{1}{2}b_i$ represents 0 and randomization $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ represents 1 for every player i. More precisely, if player i in U_k has switched at stage t_{ik}, which corresponds to the $\tilde{t}_{ik}$-th stage within U_k, then he must play the sequence of randomizations in I_k that corresponds to $\tilde{t}_{ik} - 1$ in binary code. Note that n stages are sufficient for reporting the binary code, since U_k consists of 2^n stages.[9]

Control-block C_k: The idea behind C_k is that players check whether opponents have played the "right" actions, and have bought information whenever they were supposed to.

Information purchase: At the beginning, every player i buys information about all stages in I_k, which enables him to identify the stages in U_k at which his opponents have switched. Additionally, player i buys information about these stages in U_k. More precisely, if player i learns that opponent j_1 has reported the switching stage t_{j_1k}, and opponent j_2 has reported t_{j_2k}, he buys stages t_{j_1k} and t_{j_2k} in U_k. If the two reported switching stages coincide, he only buys this stage once. (Note that each player i knows the actions that were played at his own switching stage t_{ik}. Namely, if t_{ik} coincides with an opponent's switching stage, then he must buy that stage. If, on the other hand, player i is the only player who switched at t_{ik}, then he knows what his opponents were supposed to play there.) Subsequently, player i buys information about the previous control-block C_{k-1} and the previous deviation-report-block D_{k-1} (if $k \geqslant 2$).

Actions: Two different scenarios may occur here. (a) If no player reported a deviation at D_{k-1} (we will describe later how deviations can be reported): If among all actions at stages t_{jk} ($j = 1, 2, 3$), C_{k-1} and D_{k-1} the sum of the number of a's and the number of $\bar{a}$'s is even, then player i equally randomizes between a_i and b_i. Otherwise, player i equally randomizes between $\bar{a}_i$ and $\bar{b}_i$. The purpose of these randomizations is to enforce opponents to buy information about this stage later. If all players follow the prescribed strategies, then the randomizations at C_k chosen by the different players should match. (b) If a player has reported a deviation at D_{k-1}: In this case player i buys information about all previous stages and will find out who has deviated from the prescribed strategy (possibly the player who wrongly reported at D_{k-1} that somebody deviated). From master-block M_{k+1} on, the deviator will be punished by his opponents. For instance, player 1 can be punished by his opponents if player 2 chooses b_2 and player 3 chooses a_3, giving player 1 a payoff of 0.[10,11]

Deviation-report-block D_k: The purpose of this block is to enable the players to report observed deviations.

Information purchase: Players buy no information.

Actions: If player i concludes, on the basis of his information about past stages, that some opponent has deviated from the prescribed strategy, then he equally randomizes between $\bar{a}_i$ and $\bar{b}_i$; otherwise, player i equally randomizes between a_i and b_i.[12]

In Miyagawa et al. (2004), the length of U_k is not fixed, but determined stochastically, with the expected length being long enough. At each period in U_k, players switch with a small probability, implying that the number of switches in U_k is determined stochastically. Their construction is such that every player is indifferent between playing the target action and playing the switching action. Either no player monitors, or all players monitor all opponents.

[9] In Ben-Porath and Kahneman (2003), information-exchange is simpler as it can be done through public announcements. They let every player announce both switching stages, together with his switching actions, and his observations during monitoring.

In Miyagawa et al. (2004), this block is not used.

[10] Note that in this construction, the stages t_{jk} for $j \neq i$ can only be bought *after* buying block I_k. If players could only buy one collection of past stages at every stage, block C_k could be modified as follows: Add one additional stage to C_k. At the first stage, player i plays $\frac{1}{2}a_i + \frac{1}{2}b_i$ and purchases all stages in I_k, C_{k-1} and D_{k-1}. At the second stage, the player buys the stages t_{jk} for all $j \neq i$, and plays as specified in Table 1.

[11] In Ben-Porath and Kahneman (2003), this block is not needed. By costlessly listening to the public announcements, players will immediately have sufficient information regarding past actions and possible deviations.

In our case, if a player reports a deviation, the players will find out who has deviated by buying all previous stages. In Ben-Porath and Kahneman (2003), the task of finding out who deviated is more complex. In general, only the deviator and the player who monitored him at the corresponding block will know who deviated. In Ben-Porath and Kahneman, both of these players will be punished during a certain number of stages.

In Miyagawa et al. (2004), a player has to state whether a given opponent has switched or not at a given previous stage. Communication is based on actions, and not on public announcements.

[12] In Ben-Porath and Kahneman (2003), this type of communication is not needed due to the use of public announcements.

In Miyagawa et al. (2004), a player checks the report of the reporting player in C_k, and approves or disapproves his answer by playing actions in D_k.

Recall that, before turning to the next master-block M_{k+1}, a new target-payoff-vector u_{k+1} will be constructed as to compensate for "contaminating" payoffs in M_{k-1}. By construction, such contaminating payoffs can only take place in the sub-blocks I_{k-1} (n stages), C_{k-1} (1 stage) and D_{k-1} (1 stage), as well as in the switching-stages in U_{k-1} (at most 3 stages). Hence, in total we have at least $2^n - 3$ stages with "correct" payoffs (all in U_{k-1}), and at most $n + 5$ stages with "wrong" payoffs in master-block M_{k-1}. As such, we can always choose δ and n large enough such that (a) any sequence of "wrong" payoffs can be compensated by a target payoff vector u_{k+1} which lies in the interior of *FIR*, and (b) punishment of deviations remains effective despite the fact that it will only start at the end of the next master-block.

3.3. How deviations can be detected

Suppose that we are in master-block M_k, and that at previous deviation-report-blocks $D_1, \ldots, D_{k-2}$ no player has reported a deviation. Then, deviations in M_k by player i can be detected in M_k or M_{k+1} by player j using one of the following detection mechanisms:

DM1. If player j in C_k buys information about the stage $\tilde{t}_{ik}$ reported in I_k by player i, and notices that player i did not switch at that stage.

DM2. If player j in C_k buys information about the stage $\tilde{t}_{j'k} \neq \tilde{t}_{ik}$ reported in I_k by player j', and notices that player i did not play the prescribed action at that stage. (Note that we need at least three players for this to work.)

DM3. If player j in C_{k+1} buys information about C_k and notices that player i has reported differently at C_k than himself. Here, by "differently" we mean that player i has chosen a duplicate action while player j has not, or vice versa.

DM4. If player j in C_{k+1} buys information about D_k, notices that player i has reported a deviation at D_k, buys all previous stages, and finds out that player i incorrectly reported a deviation.

3.4. Why it is not profitable to deviate

We shall now explain why in the strategy profile above it is not profitable for any player i to deviate at any stage by choosing a different action and/or a different information purchase. We show for any deviation in M_k, which without punishment could be profitable, that such deviation will be detected, either in M_k or in M_{k+1}, with a positive probability which does not depend on δ. Therefore, every such deviation, when detected, will be reported in D_k or D_{k+1}. Hence, punishment of such deviations, when detected, will start no later than at M_{k+3}. Therefore, by choosing δ sufficiently close to 1, the punishments will be severe enough to render such deviations unprofitable.

Assume that player i is the only player to deviate from the strategy profile described above. We distinguish the following four exhaustive and mutually exclusive cases.

Case 1. Assume that player i's first deviation is at U_k. There are three possibilities:

Case 1.1. Assume that the first deviation is to choose a different probability distribution over possible switching-stages. In fact, it does not matter how player i chooses the probability distribution over possible switching-stages, since the modification of the target-payoff at the master-block M_{k+2} makes player i indifferent between any switching-stage that can be chosen.

Case 1.2. Assume that the first deviation is to choose a different probability distribution over actions. (a) If player i does not switch at all at U_k, this will eventually be detected with probability 1 by DM1, since at I_k player i must report a switching-stage. (b) If player i switches in U_k, but chooses a different probability distribution over the switching-actions at the switching-stage, then this is not profitable by the target-payoff modification at the master block M_{k+2}. (c) If player i switches in U_k but chooses a different action at a non-switching stage, then either this stage or the switching-stage will not be reported at I_k. Let $\hat{t}_{ik}$ denote this stage. Then, with probability at least $c/(2^n - 1)(2^n + c)$ this stage $\hat{t}_{ik}$ coincides with the switching stage $t_{j'k}$ of some third player j' and will therefore be bought by player j at C_k. Therefore, player i will be "caught" with probability at least $c/(2^n - 1)(2^n + c)$ by DM2.

Case 1.3. Assume that the first deviation is to buy information at some stage in U_k. The only useful insights that player i can obtain by buying information about past stages is to learn whether his opponents have already switched at U_k. If he learns that both opponents have already switched, he could deviate at all future non-switching-stages in U_k without running the risk of being detected by his opponents, since only reported switching-stages will be bought (at C_k). Suppose that player i at stage t_2 in U_k is still not certain whether opponent j has switched in U_k, and buys

some stage $t_1 < t_2$ in U_k. Hence, he has a cost of $(1 - \delta)\delta^{t_2-1}c$. Moreover, as player i is uncertain about whether player j switched at t_1 or will switch at the last stage in U_k, the conditional probability that player j has switched at stage t_1 is at most

$$\frac{c/(2^n - 1)(2^n + c)}{c/(2^n - 1)(2^n + c) + 2^n/(2^n + c)} = \frac{c}{c + 2^n(2^n - 1)}.$$

(Note that the probability of switching at stage t_1 in U_k is initially equal to $c/(2^n - 1)(2^n + c)$, while the last stage is chosen with probability $2^n/(2^n + c)$.) If player j indeed switched at t_1, then player i does not have to buy this switching stage anymore at C_k, and in the best-case scenario player i would moreover know that all opponents have switched before stage t_2, in which case he could safely deviate at the remaining stages in U_k. So, the total gain would be at most

$$(1 - \delta)\left(\delta^{z_k+2^n+n-1}c + \sum_{t=t_2}^{z_k+2^n-1} \delta^{t-1} \cdot 1\right),$$

where z_k denotes the first stage in U_k. (Notice that $(1 - \delta)\delta^{z_k+2^n+n-1}c$ is the cost which player i would have to incur if he would buy stage t_1 at C_k. Note also that the highest possible gain at each stage is at most 1, as can be seen from the payoff matrices.) Therefore, the expected total gain from buying stage t_1 at stage t_2 is at most

$$\frac{c}{c + 2^n(2^n - 1)} \cdot (1 - \delta)\left(\delta^{z_k+2^n+n-1}c + \sum_{t=t_2}^{z_k+2^n-1} \delta^{t-1} \cdot 1\right).$$

Since

$$\frac{c}{c + 2^n(2^n - 1)} \cdot (1 - \delta)\left(\delta^{z_k+2^n+n-1}c + \sum_{t=t_2}^{z_k+2^n-1} \delta^{t-1} \cdot 1\right)$$

$$\leqslant \frac{c}{c + 2^n(2^n - 1)} \cdot (1 - \delta)\delta^{t_2-1}\left(c + 2^n\right)$$

$$\leqslant (1 - \delta)\delta^{t_2-1}c$$

such deviation can never be profitable in expectation, as $(1 - \delta)\delta^{t_2-1}c$ is the discounted cost of buying information at stage t_2 about stage t_1.

Case 2. Assume that player i's first deviation is at I_k. This implies that player i has switched exactly once, at stage t_{ik}, in U_k. If player i chooses a different probability distribution over the prescribed actions, it would not be profitable by the target-payoff modification at the master block M_{k+2}. If player i reports a stage which is different from t_{ik}, this will be detected by DM1. Therefore, any information that player i would buy at I_k cannot be used in I_k since it is in player i's best interest to report stage t_{ik} here. Hence, buying such information is best delayed until the beginning of block C_k.

Case 3. Assume that player i's first deviation is at C_k. First of all, choosing a different probability distribution over the prescribed actions is not profitable by the target-payoff modification at the master block M_{k+2}. If he deviates by choosing a non-prescribed action, then, by definition, it will be detected by DM3. Finally, if player i buys less information at C_k, then, with probability $1/2$, player i will be wrong about the parity of the sum of the number of a's and the number of $\bar{a}$'s at the reported switching stages and blocks C_{k-1}, D_{k-1}. Hence, with probability $1/2$ he reports differently than player j at C_k. Consequently, with probability $1/2$ he chooses a non-prescribed action which will be detected by DM3.

Case 4. Assume that player i's first deviation is at D_k. Since we assume that player i is the only player who deviates, player i is supposed to play $\frac{1}{2}a_i + \frac{1}{2}b_i$ here. First of all, it does not matter how player i chooses the probability distribution over a_i and b_i, since the modification of the target-payoff at master-block M_{k+2} makes player i indifferent between them. If player i chooses a duplicate action, thereby wrongly reporting a deviation, this will be detected in C_{k+1} by DM4. Obviously, buying information at D_k is useless.

4. Formal proof of main result

4.1. Outline

Consider a repeated game satisfying full dimensionality. Suppose there are $m \geqslant 3$ players and every player has at least four actions. Take an arbitrary payoff vector $u \in \mathrm{int}(FIR)$. Our goal is to find a lower bound $\delta(u) \in (0, 1)$ for the discount factor, and for every $\delta \in (\delta(u), 1)$ to construct a Nash equilibrium of the repeated game with expected δ-discounted payoff vector equal to u. For a fixed u the outline of the construction is as follows:

In Section 4.2 we first show how to choose $\delta(u)$. Take an arbitrary discount factor $\delta \in (\delta(u), 1)$. In Section 4.3 we construct a strategy profile. In Section 4.4 we show that it has expected δ-discounted payoff u. In Section 4.5, finally, we prove that it is an equilibrium in the δ-discounted repeated game. This will complete the proof of Theorem 3.1.

4.2. Choice of $\delta(u)$

Recall that we assume that the dimension of FIR is equal to m. Therefore, there exists some $\alpha > 0$ such that $B_{4\alpha}(u) \subseteq FIR$, where $B_{4\alpha}(u)$ denotes the open ball with center u and radius 4α with respect to the maximum norm.

Let $R := \{(v_i(a))_{i \in I} \mid a \in \times_{i \in I} A_i\}$ be the set of stage-payoff vectors induced by pure action profiles. Hence, the set of feasible payoff vectors is the convex hull of R. Let $r^* := 2 \max_{r \in R} \|r\|$, where $\| \cdot \|$ denotes the maximum norm. Therefore, the maximum variation in payoff at a certain stage can never exceed r^*. Given α, choose a large enough N of the form $N = 2^n + n + 2$, with $n \in \mathbb{N}$, such that the following two properties hold:

(1) Every payoff-vector $r \in FIR$ can be approximated sufficiently closely by the average payoff of a sequence of N pure action profiles. Formally, for every $r \in FIR$ there exists a sequence $r_1, r_2, \ldots, r_N$ in R such that

$$\left\| r - \frac{r_1 + r_2 + \cdots + r_N}{N} \right\| \leqslant \frac{\alpha}{12}.$$

(2) Variations in stage-payoffs during at most $\log_2 N + m + 2$ stages within a block of N consecutive stages have little influence on the average payoff. Formally,

$$\frac{\lceil \log_2 N \rceil + m + 2}{N} r^* \leqslant \frac{\alpha}{12},$$

where $\lceil x \rceil$ denotes the lowest integer above or equal to x.

Subsequently, we choose $\delta(u) \in (0, 1)$ close enough to 1 such that for every $\delta \in (\delta(u), 1)$ the following properties hold:

(3) For every sequence $r_1, r_2, \ldots, r_N$ in R, it holds that the δ-discounted sum is close enough to the average. Formally,

$$\left\| \frac{\sum_{t=1}^{N} \delta^{t-1} r_t}{\sum_{t=1}^{N} \delta^{t-1}} - \frac{r_1 + r_2 + \cdots + r_N}{N} \right\| \leqslant \frac{\alpha}{12}.$$

(4) The δ-discounted influence of variations during at most $\log_2 N + m + 2$ stages within a block of N consecutive stages is close enough to the average influence. Formally,

$$\left| \frac{\sum_{t=1}^{\lceil \log_2 N \rceil + m + 2} \delta^{t-1}}{\sum_{t=1}^{N} \delta^{t-1}} r^* - \frac{\lceil \log_2 N \rceil + m + 2}{N} r^* \right| \leqslant \frac{\alpha}{12}.$$

(5) Even after two blocks of N stages, the payoffs are not discounted "too much." Formally,

$$\delta^{2N} > \frac{1}{3}.$$

(6) The maximum variation in δ-discounted payoff during three blocks of N stages is limited. Formally,

$$\sum_{t=1}^{3N} \delta^{t-1} r^* \leqslant \frac{c}{N(N+c)} \sum_{t=3N+1}^{\infty} \delta^{t-1} \alpha.$$

4.3. Strategy profile

For a given u, let $N = 2^n + n + 2$ and $\delta(u)$ be chosen as specified above. Take an arbitrary $\delta \in (\delta(u), 1)$. By properties (1) and (3) above, one can construct for every $r \in FIR$ a sequence $\xi(r) := (\xi_1(r), \xi_2(r), \dots, \xi_N(r))$ in R with

$$\left\| \frac{\sum_{t=1}^{N} \delta^{t-1} \xi_t(r)}{\sum_{t=1}^{N} \delta^{t-1}} - r \right\| \leqslant \frac{\alpha}{6}.$$

In fact, we only need $\xi_t(r)$ for $t \leqslant 2^n$, but we include N stages for technical reasons. For every $r \in FIR$, define

$$\varphi(r) := \frac{\sum_{t=1}^{N} \delta^{t-1} \xi_t(r)}{\sum_{t=1}^{N} \delta^{t-1}}.$$

Similarly to our example in Section 3, we build a sequence of blocks $(U_1, I_1, C_1, D_1, U_2, \dots)$ where each utility block U_k has 2^n stages, each information-exchange-block I_k has n stages, each control block C_k has 1 stage and each deviation-report block D_k has 1 stage. Hence, every master-block M_k (consisting of U_k, I_k, C_k and D_k) has length N. The purpose of each of these blocks is the same as in the example. We define a strategy profile as follows:

In block U_1, the target-payoff vector u_1 is chosen equal to u. At the beginning of U_1, every player i randomly chooses a switching-stage t_{i1} according to the probability distribution specified in Table 1. At all stages t of U_1, player i chooses the pure action corresponding to $\xi_t(u_1)$, except for the switching stage t_{i1} at which he randomizes equally between the two non-prescribed actions with lowest indices, say $\bar{a}_i$ and $\bar{b}_i$.

For the blocks I_1, C_1 and D_1, we may take for each player i these two actions $\bar{a}_i$ and $\bar{b}_i$ together with two other actions, say a_i and b_i, for communication. The behavior of the players in these blocks is then identical to Table 1.

In block U_2 the target-payoff vector u_2 is still equal to u, and the behavior of the players in master-block M_2 is the same as in M_1.

If all players follow the prescribed strategy, every player knows after C_2 the realized actions and payoffs in M_1. Let $(r_t)_{t \in M_1}$ denote the sequence of realized[13] payoff vectors in M_1. Given these realized payoffs, choose the target-payoff vector u_3 for U_3 such that

$$\sum_{t \in M_1} \delta^{t-1}(r_t - u_1) + \sum_{t \in M_3} \delta^{t-1} u_3 = \sum_{t \in M_3} \delta^{t-1} u.$$

This way, u_3 compensates for the difference between the target payoff u_1 and the realized payoff in block M_1. The equation above gives

$$u_3 = \frac{\sum_{t \in M_1} \delta^{t-1}(u_1 - r_t) + \sum_{t \in M_3} \delta^{t-1} u}{\sum_{t \in M_3} \delta^{t-1}}.$$

At the tth stage of U_3, every player i chooses the pure action corresponding to $\xi_t(u_3)$, except when t is i's switching stage. The remainder of block M_3 is the same as in M_1 and M_2.

In general, for every $k \geqslant 3$, let $(r_t)_{t \in M_{k-2}}$ denote the sequence of realized payoff vectors in M_{k-2}. We choose the target-payoff vector u_k for U_k such that

$$\sum_{t \in M_{k-2}} \delta^{t-1}(r_t - u_{k-2}) + \sum_{t \in M_k} \delta^{t-1} u_k = \sum_{t \in M_k} \delta^{t-1} u, \tag{4.1}$$

and hence

$$u_k = \frac{\sum_{t \in M_{k-2}} \delta^{t-1}(u_{k-2} - r_t) + \sum_{t \in M_k} \delta^{t-1} u}{\sum_{t \in M_k} \delta^{t-1}}.$$

The remainder of block M_k is as usual.

In order to show that these strategies are well-defined, we prove that $u_k \in \text{int}(FIR)$ for all k. To this purpose, we shall show that $\|u_k - u\| < \alpha$ for all k, implying that $u_k \in B_\alpha(u) \subseteq \text{int}(FIR)$ for all k.

---

[13] More precisely, each player has a belief about the payoffs in U_1. If no player deviates, then all these beliefs will coincide with r_t.

For $k = 1$ and $k = 2$, the statement holds trivially since $u_1 = u_2 = u$. By equation (4.1), we have for every $k \geqslant 3$ that

$$
\begin{aligned}
\|u_k - u\| &= \left\| \frac{\sum_{t \in M_{k-2}} \delta^{t-1}(u_{k-2} - r_t)}{\sum_{t \in M_k} \delta^{t-1}} \right\| \\
&= \frac{1}{\delta^{2N}} \left\| \frac{\sum_{t \in M_{k-2}} \delta^{t-1}(u_{k-2} - r_t)}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| \\
&= \frac{1}{\delta^{2N}} \left\| u_{k-2} - \frac{\sum_{t \in M_{k-2}} \delta^{t-1} r_t}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| \\
&\leqslant \frac{1}{\delta^{2N}} \left(\|u_{k-2} - \varphi(u_{k-2})\| + \left\| \varphi(u_{k-2}) - \frac{\sum_{t \in M_{k-2}} \delta^{t-1} r_t}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| \right).
\end{aligned}
\tag{4.2}
$$

By definition of $\varphi(u_{k-2})$,

$$
\left\| \varphi(u_{k-2}) - u_{k-2} \right\| \leqslant \frac{\alpha}{6}.
\tag{4.3}
$$

Again, by definition of $\varphi(u_{k-2})$,

$$
\begin{aligned}
\left\| \varphi(u_{k-2}) - \frac{\sum_{t \in M_{k-2}} \delta^{t-1} r_t}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| &= \left\| \frac{\delta^{(k-3)N} \sum_{t=1}^{N} \delta^{t-1} \xi_t(u_{k-2})}{\delta^{(k-3)N} \sum_{t=1}^{N} \delta^{t-1}} - \frac{\sum_{t \in M_{k-2}} \delta^{t-1} r_t}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| \\
&= \left\| \frac{\sum_{t \in M_{k-2}} \delta^{t-1}(\xi_t(u_{k-2}) - r_t)}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| \\
&\leqslant \frac{\sum_{t=1}^{m+n+2} \delta^{t-1} r^*}{\sum_{t=1}^{N} \delta^{t-1}},
\end{aligned}
$$

since, by construction, there are at most $m + n + 2$ stages in M_{k-2} where the realized payoff r_t is different from the target payoff $\xi_t(u_{k-2})$. Since $m + n + 2 \leqslant \log_2 N + m + 2$, we have by properties (2) and (4) that

$$
\frac{\sum_{t=1}^{m+n+2} \delta^{t-1} r^*}{\sum_{t=1}^{N} \delta^{t-1}} \leqslant \frac{\sum_{t=1}^{\lceil \log_2 N \rceil + m + 2} \delta^{t-1} r^*}{\sum_{t=1}^{N} \delta^{t-1}} \leqslant \frac{\alpha}{6},
$$

and hence

$$
\left\| \varphi(u_{k-2}) - \frac{\sum_{t \in M_{k-2}} \delta^{t-1} r_t}{\sum_{t \in M_{k-2}} \delta^{t-1}} \right\| \leqslant \frac{\alpha}{6}.
\tag{4.4}
$$

By inequalities (4.2), (4.3) and (4.4) it follows that

$$
\|u_k - u\| \leqslant \frac{1}{\delta^{2N}} \left(\frac{\alpha}{6} + \frac{\alpha}{6} \right) < \alpha,
$$

where the last inequality follows from property (5). We thus have shown that $u_k \in B_\alpha(u) \subseteq \text{int}(FIR)$ for all k, and therefore the strategy profile is well-defined.

4.4. Expected δ-discounted payoff is equal to u

Take a discount factor $\delta \in (\delta(u), 1)$ and consider the corresponding strategy profile as defined in Section 4.3. From (4.1) it follows that

$$
\sum_{k=1,\, k\,\text{odd}}^{\infty} \sum_{t \in M_k} \delta^{t-1} r_t = \sum_{k=3,\, k\,\text{odd}}^{\infty} \sum_{t \in M_{k-2}} \delta^{t-1} r_t
$$

$$= \sum_{k=3, k \text{ odd}}^{\infty} \left(\sum_{t \in M_{k-2}} \delta^{t-1} u_{k-2} - \sum_{t \in M_k} \delta^{t-1} u_k + \sum_{t \in M_k} \delta^{t-1} u \right)$$

$$= \sum_{t \in M_1} \delta^{t-1} u_1 + \sum_{k=3, k \text{ odd}}^{\infty} \sum_{t \in M_k} \delta^{t-1} u$$

$$= \sum_{k=1, k \text{ odd}}^{\infty} \sum_{t \in M_k} \delta^{t-1} u.$$

Similarly, one can show that

$$\sum_{k=2, k \text{ even}}^{\infty} \sum_{t \in M_k} \delta^{t-1} r_t = \sum_{k=2, k \text{ even}}^{\infty} \sum_{t \in M_k} \delta^{t-1} u.$$

Together, we obtain

$$\sum_{k=1}^{\infty} \sum_{t \in M_k} \delta^{t-1} r_t = \sum_{k=1}^{\infty} \sum_{t \in M_k} \delta^{t-1} u,$$

and hence

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} r_t = u,$$

which means that, with probability 1, the total δ-discounted realized payoff is exactly u. This implies in particular that the expected δ-discounted payoff is equal to u, as desired.

4.5. The strategy-profile is an equilibrium in the δ-discounted repeated game

Consider, again, a discount-factor $\delta \in (\delta(u), 1)$ and the corresponding strategy profile. Suppose that player i would deviate for the first time at master-block M_k, and that no other player has deviated before. In Section 3.3 we have described how player i's opponents can detect such a deviation by means of the detection mechanisms DM1 until DM4. Moreover, in Section 3.4 we have shown that any deviation that could possibly lead to an improvement will be detected with probability at least $c/(2^n - 1)(2^n + c)$. By construction, if the deviation is detected, it will be reported no later than in D_{k+1}, and hence will be noticed by the other players no later than in C_{k+2}. Therefore, punishment will start no later than in M_{k+3}. These arguments are also valid for the general strategy profile we consider here. We will now show that punishments are severe enough to render any deviation unprofitable.

By deviating in block M_k, the maximum δ-discounted gain that player i can achieve in blocks M_k, M_{k+1} and M_{k+2} is

$$\sum_{t \in M_k \cup M_{k+1} \cup M_{k+2}} \delta^{t-1} r^*.$$

If the deviation will be detected, his loss in any master-block M_l ($l \geqslant k + 3$) will be at least

$$\sum_{t \in M_l} \delta^{t-1} \alpha.$$

For if player i did not deviate, the target-payoff vector u_l would, by construction, be in $B_\alpha(u)$. On the other hand, α was chosen such that $B_{2a}(u) \subseteq FIR$, which implies that i's punishment payoff is not even in $B_{2a}(u)$. Therefore, the distance between i's punishment payoff and i's payoff in u_l is at least α. Since i's deviation will be detected at least with probability $c/(2^n - 1)(2^n + c)$, player i's expected gain from deviating at M_k will be at most

$$\sum_{t \in M_k \cup M_{k+1} \cup M_{k+2}} \delta^{t-1} r^* - \frac{c}{(2^n - 1)(2^n + c)} \sum_{l \geqslant k+3} \sum_{t \in M_l} \delta^{t-1} \alpha$$

$$= \delta^{(k-1)N} \left(\sum_{t=1}^{3N} \delta^{t-1} r^* - \frac{c}{(2^n - 1)(2^n + c)} \sum_{t=3N+1}^{\infty} \delta^{t-1} \alpha \right)$$

$$\leqslant \delta^{(k-1)N} \left(\sum_{t=1}^{3N} \delta^{t-1} r^* - \frac{c}{N(N+c)} \sum_{t=3N+1}^{\infty} \delta^{t-1} \alpha \right)$$

$$\leqslant 0,$$

where the last inequality follows from property (6). Hence, no deviation is profitable, which completes the proof.

5. Sequential equilibrium

In this section we prove that our main result does not only hold for Nash equilibrium, but also for sequential equilibrium. A strategy profile is called a *sequential equilibrium* if for every player i and *every* possible history for player i, his continuation strategy is optimal against the opponents' expected continuation strategies. Below we will construct, for every payoff-vector u in int(FIR) and every discount factor δ close enough to 1, a sequential equilibrium σ that induces the payoff-vector u. Moreover, the total cost for buying information in σ will be close to 0 (in fact, tends to 0 if δ approaches 1).

Note that the strategy profile as constructed in Section 3 will in general not be a sequential equilibrium. The problem arises after a player detects a deviation. In that case, the continuation strategy of a player will in general not be optimal against the opponents' expected continuation strategies. Often, it would not be in a player's interest to report a deviation, because according to the strategy profile he would have to buy all previous stages in this case, and punish the deviating player afterwards. These two activities would in general hurt the player.

In order to overcome these problems, we now make sure that it becomes optimal for a player to report an opponent's deviation whenever he detects one. To this purpose, we make the following adjustments when a player reports a deviation:

1. Players do not have to buy previous stages to find out the true identity of the deviator, because the deviator will be named by one of the players, say player j, who reported a deviation. A target-payoff modification will assure that player j is indifferent between any opponent he can name, so he is trustworthy.[14]

2. Punishments now take place during finitely many stages only. When necessary, the reporting player is compensated (or charged) for possible losses (or gains) caused by the punishments by means of a target-payoff modification.

5.1. Outline of the plan of action

A plan of action is a description of a player's behavior as long as he has not deviated himself from this plan. We will now describe the plan of action τ_i for every player i. Later, in Section 5.4, we will extend the plan of action τ_i to a complete strategy σ_i, which also describes what player i will do in case he has deviated from τ_i. In the previous sections it was sufficient to describe a plan of action for every player since the concept of Nash equilibrium, in contrast with sequential equilibrium, does not require a player i to choose optimally after he has deviated himself from τ_i.

Just as in Section 3 play is divided into different master-blocks, in this case regular master-blocks and punishment master-blocks. Regular master-blocks M_k are defined similarly to Section 3, except for some slight modifications that occur after a punishment. These changes will be described in Section 5.3. Punishment master-blocks $P[j]$ will be really new, and are meant to punish a player during a finite number of stages. Section 5.2 provides a detailed description of $P[j]$. Here, j denotes the player who can name a deviator. The transition between these blocks proceeds according to the following inductive scheme:

Play starts with a regular master-block M_1.

After regular master-block M_k, the new master-block is determined as follows:

Let $P[j']$ be the last punishment master-block before M_k. If no punishment has yet taken place, let j' be player 1.

---

[14] He may not be trustworthy if he is the only deviator, because in this case he cannot name an opponent who deviated.

∗ If no player other than j' reported a deviation in D_{k-1} (which is bought at M_k), we enter a regular master-block M_{k+1}.[15]

∗ Otherwise, we first enter a punishment master-block $P[j]$, after which we go to a regular master-block M_{k+1}. The reporting player j in $P[j]$ is determined as follows: Let j_1 and j_2 be the players with the lowest respectively highest index not equal to j' who reported a deviation in D_{k-1} (possibly $j_1 = j_2$). If the sum of the number of a's and $\bar{a}$'s at D_{k-1} is even, then the reporting player j will be j_1. If the sum is odd, then j will be j_2.[16]

5.2. Plan of action for the punishment master-block

As we discussed above, the major difference with Section 3 lies in the construction of the punishments. In a punishment master-block, following a regular master-block M_k and initiated by a reported deviation in D_{k-1}, there is a special role for a player j who can name a deviating player. The corresponding punishment master-block will be called $P[j]$. The play in this phase will consist of four blocks, $R^*[j]$, C^*, P^* and C^{**}. Table 2 provides a summary of the players' behavior in $P[j]$.

1. Block $R^*[j]$: The purpose of this block is to enable player j to name a player j^* as the deviator. Player j must choose randomly amongst the players whose deviations he has detected (including players who deviated at D_{k-1}, that is, players who should have reported a deviation in D_{k-1} but omitted to do so, and players who falsely reported a deviation at D_{k-1}).[17]

This block has length $\lceil \log_2 m \rceil$, where m is the number of players.

Information purchase: During block $R^*[j]$ no player buys any information.

Actions: Player j will name in binary code the identity of a player as the deviator. Recall that we use randomization $\frac{1}{2}a_j + \frac{1}{2}b_j$ for digit 0 and $\frac{1}{2}\bar{a}_j + \frac{1}{2}\bar{b}_j$ for digit 1. Hence, $\lceil \log_2 m \rceil$ stages are sufficient for reporting such binary code. All other players i choose $\frac{1}{2}a_i + \frac{1}{2}b_i$.

2. Block C^*: The purpose of this block is to guarantee that all players buy all stages in $R^*[j]$, so that all players will know the player j^* that has been reported by player j as the deviator. This block consists of 1 stage.

Information purchase: Each player buys all stages in $R^*[j]$. Now all players will know j^* (if player j reported a number that does not correspond to a player, then let $j^* = j$).

Table 2
Plans of action in punishment master-block $P[j]$

	Actions	Information purchase
$R^*[j]$ ($\lceil \log_2 m \rceil$ stages)	∗ Player j reports identity of deviator j^* in binary code ∗ Every non-reporting player i plays $\frac{1}{2}a_i + \frac{1}{2}b_i$	none
C^* (1 stage)	∗ Every player i plays $\frac{1}{2}a_i + \frac{1}{2}b_i$ if the sum of the number of a's and $\bar{a}$'s at $R^*[j]$ is even ∗ Plays $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ otherwise	$R^*[j]$
P^* (finitely many stages)	∗ All punishing players $i \neq j^*$ play min-max action against j^* (with small perturbation) ∗ Punished player j^* plays best reply against opponents' actions at every stage	none
C^{**} (1 stage)	∗ Every player i plays $\frac{1}{2}a_i + \frac{1}{2}b_i$ if the sum of the number of a's and $\bar{a}$'s at P^* is even ∗ Plays $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ otherwise	P^*

---

[15] By ignoring a possible report by j' alone, we avoid the situation where j' would have to keep on reporting deviations, which would imply that he could no longer be compensated for carrying out so many punishments. By our construction, the next player who can report the deviator must be different from j', so that after the punishment we can apply a different kind of target-payoff modification for j'. This then enables us to position the new target-payoff for j' such that he could be compensated next time.

[16] This lottery between j_1 and j_2 makes sure that it cannot happen that a player deviates in U_{k-1} while knowing that, with probability 1, he will be able to name somebody else as the deviator, and thereby avoiding being punished. Even though $j_1 = j_2$ is possible, you can never know this in advance.

[17] Note that in a plan of action τ_j player j cannot be a deviator himself.

Actions: Every player i plays $\frac{1}{2}a_i + \frac{1}{2}b_i$ if the sum of the number of a's and the number of $\bar{a}$'s at $R^*[j]$ is even, and plays $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ otherwise.

3. Punishment-block P^*: The purpose of this block is to sufficiently punish player j^*. The length of this block can be chosen dependent of the length of a regular master-block, but independent of δ (given δ is large enough).

Information purchase: No player buys any information.

Actions: Let α_{-j^*} be the mixed action profile for j^*'s opponents such that

$$d_{j^*} = \max_{a_{j^*} \in A_{j^*}} v_{j^*}(a_{j^*}, \alpha_{-j^*}).$$

Hence, α_{-j^*} is the min-max mixed action-profile against j^*. At every stage in P^*, every punishing player $i \neq j^*$ plays with high probability his mixed action α_i in α_{-j^*} and with low probability any other action in A_i. As such, every action in A_i will be played with positive probability at every stage in P^*. The probability of playing α_i must be chosen (1) high enough so that player j^*'s expected payoff at every stage does not exceed d_{j^*} by much, and (2) low enough so that at P^* there will be sufficient uncertainty about the actions, and therefore it will be in the players' interest to buy every stage of P^* at the following block C^{**}.[18] The punished player j^* plays at every stage a best reply against the mixed action-profile of his opponents.

4. Block C^{**}: The purpose of this block is to guarantee that all players buy all stages in P^*. This is important for the target-payoff modification later on. This block consists of 1 stage.

Information purchase: Each player buys all stages in P^*.

Actions: Every player i plays $\frac{1}{2}a_i + \frac{1}{2}b_i$ if the sum of the number of a's and the number of $\bar{a}$'s at P^* is even, and plays $\frac{1}{2}\bar{a}_i + \frac{1}{2}\bar{b}_i$ otherwise.

5.3. Plan of action for a regular master-block after a punishment master-block

As we know from Section 5.1, after a punishment master-block $P[j]$, following a regular master-block M_k and initiated by a reported deviation in D_{k-1}, we go back to a regular master-block M_{k+1}. This regular master-block is defined as in Section 3, with the only exceptions that now (1) at C_{k+1} players will also buy information about the blocks C^* and C^{**} in $P[j]$, and the report at C_{k+1}, should also depend on the observed actions at these two blocks, and (2) the target-payoff has to be constructed in a different way.

The idea behind the target-payoff modification is as follows. We will choose the new target-payoff vector such that it lies in $B_{2\alpha}(u) \subseteq \text{int}(FIR)$. Each player's target-payoff should take the realized actions at $R^*[j]$ and P^* into account, so that each player is indifferent between the actions at $R^*[j]$ and P^* on which he was supposed to randomize.[19] Also, the reporting player j in $R^*[j]$ should be made indifferent between all players he can report, as far as the payoffs is $R^*[j]$ are concerned. Further, the target-payoff for the player who was punished, say player j^*, is chosen in such a way that punishment remains effective. For all players except j and j^* the target-payoff is chosen such that, if they become the reporting player next time, they can be compensated (this means that the target-payoff should stay away from the boundary of FIR). Moreover, their target-payoff should be independent of j and j^*. For the target-payoff of player j we distinguish two cases.

Case 1. Assume that player j was the only player, except for possibly j', who reported a deviation at D_{k-1} (recall that $P[j']$ was the last punishment-master-block; if no punishment has yet taken place then $j' = 1$). In this case, the target-payoff for player j is modified as to compensate (or charge) him for possible losses (or gains) in payoffs during $R^*[j]$ and P^*.

Case 2. Assume that there is a player other than j and j' who reported a deviation at D_{k-1}. In this case, the target-payoff for player j is modified just as for the other players, except j^*.

---

[18] This idea is rather standard.

[19] Note that the actions at the other two blocks C^* and C^{**} will be bought at C_{k+1}, and the target-payoff modification at the following regular master-block M_{k+2} will take these actions into account.

5.4. Construction of the strategy profile

The previous three subsections together describe a plan of action τ_i for player i, but not a complete strategy yet. The remaining problem is that we did not describe what player i should do in case he has not acted consistently with τ_i in the past. We will now completely define player i's strategy σ_i.

We distinguish the following cases:

Case 1. Assume that during h_i player i has bought less or more information than prescribed by τ_i. Then, player i should calculate, for every opponent j, the expected continuation strategy under the assumption that j has bought exactly the information as prescribed by τ_j (meaning that j is not in Case 1, but in Case 2 below). Moreover, player i should play a best response against these expected continuation strategies.

Case 2. Assume that during h_i player i has bought exactly the information as prescribed by τ_i.

In this case, player i knows exactly all switching stages by his opponents, and therefore knows the stages that other players were supposed to buy.

Suppose the game is in a certain master-block, and that h_i is the corresponding history for player i. We say that a block is *reported* in h_i if in h_i the players had the opportunity to report any deviation in this block. More precisely, let D_k be the last deviation-report-block in h_i. Then, all blocks before D_k, except C_k, are reported in h_i.

Case 2.1. Assume that in all non-reported blocks player i has acted consistently with τ_i.[20] Then his behavior $\sigma_i(h_i)$ at the current stage is given by τ_i.

Case 2.2. Assume that player i has played inconsistently with τ_i in some non-reported block.[21]

Case 2.2.1. Assume that play is at C_k in regular master-block M_k, or at C^*, P^* or C^{**} in punishment master-block $P[j]$. Then, player i continues with the master-block according to τ_i (as if he had not deviated).

Case 2.2.2. Assume that play is at $R^*[j]$ in punishment master-block $P[j]$, following a regular master-block M_k, and initiated by a reported deviation in D_{k-1}. If $i \neq j$, then player i continues according to τ_i (as if he had not deviated). If $i = j$ (so i is the reporting player) then let I^* be the set of opponents whom player i can still report given his previous actions at $R^*[j]$. If I^* is empty, then player i plays an arbitrary action (he cannot avoid getting detected anyhow). Otherwise, player i randomly chooses an opponent in I^* and plays the corresponding action.

Case 2.2.3. Assume that play is at D_k in a regular master-block M_k. If player i believes that his deviation has been detected by an opponent other than j'[22] (recall that $P[j']$ was the last punishment-master-block; if no punishment has yet taken place, then $j' = 1$), or has detected a deviation by an opponent, he reports a deviation at this stage. Otherwise, he does not report a deviation.

Case 2.2.4. Assume that play is at U_k or I_k in a regular master-block M_k. Conditionally on player i behaving according to τ_i from C_k on (as if he had not deviated), he can calculate by backward induction an optimal substrategy for the remaining stages until C_k. Player i will act according to this optimal substrategy.

5.5. Strategy profile is a sequential equilibrium

It should be clear that the strategy profile constructed above induces the desired discounted expected payoff u, as it coincides with the strategy profile in Section 3 as long as no deviation has taken place. We will now show that the strategy profile constructed above constitutes a sequential equilibrium. That is, for every history $h_i \in H_i$, we must show that player i's continuation strategy $\sigma_i[h_i]$ is optimal against the opponents' expected continuation strategies. Suppose that player i considers a deviation after h_i. We will show that such a deviation cannot be profitable. We distinguish six cases here:

---

[20] Note that the construction of τ_i directly extends to this situation.

[21] The idea in this case is to let player i return to τ_i as soon as he can do so without running additional risk of getting caught. In the remaining part of Case 2.2, we describe player i's optimal substrategy until he can return to τ_i. Obviously, the existence of such an optimal substrategy on finitely many stages is guaranteed. In fact, we do more than necessary since we would only need to show that such a substrategy allows player i to report in D_k a deviation he detected (see Case 2.2.3). Note that it was essential in our construction that each player takes part in the detection process—this is why we needed at least three players.

[22] Recall that player i knows exactly all switching stages by his opponents, and therefore knows the stages that other players were supposed to buy. This implies that player i knows who is supposed to have detected his deviation.

Case A. Assume that player i considers a deviation in Case 1 of Section 5.4. In this case, a possible belief for player i is to believe that every opponent j has bought exactly the information as prescribed in τ_j. Hence, player i is supposed to act optimally given this belief, and can therefore not profitably deviate after h_i.

Case B. Assume that player i considers a deviation in Case 2.1.

Case B.1. Assume that player i considers a deviation in a regular master-block M_k.

Case B.1.1. Assume that the deviation is planned for block U_k, I_k or C_k. Then player i believes that this deviation will be detected by the opponents in exactly the same way as in Section 3. The eventual punishment will make sure that such deviation is unprofitable.[23]

Case B.1.2. Assume that the deviation is planned for block D_k. Recall that j' is the last player who reported a deviation before M_k which actually led to a punishment; if no such report has been made before M_k, then j' is chosen equal to player 1. If $i = j'$ then, by construction, i has no influence whether or not a punishment will follow, and the target-payoff modification will assure that i is indifferent between any action he can take. So, we will assume in the following subcases that $i \neq j'$.

Let $\tilde{I}$ be the set of opponents of player i, except j', who, according to player i, should report a deviation in D_k.

Case B.1.2.1. Assume that $\tilde{I}$ is empty. By reporting a deviation, player i believes that he will be the only reporting player, except j', and will therefore believe that he will be compensated by the target-payoff modification (cf. Section 5.3, Case 1). This would give him the same expected discounted payoff as he would get by not reporting. Player i is therefore indifferent between reporting and not reporting.

Case B.1.2.2. Assume that $\tilde{I}$ is not empty. We distinguish two cases.

Case B.1.2.2.1. Assume that $\tilde{I}$ contains only one player, say j. Then, player i knows that there has been a deviation. If player i deviates by not reporting this deviation, then he believes that j will be the only player other than j' who reports a deviation in D_k, and that he will name, with a positive probability, player i as the deviator in $R^*[j]$. Hence, player i will be punished, with positive probability, by not reporting this deviation.

Case B.1.2.2.2. Assume that $\tilde{I}$ contains more than one player. Then, it does not make a difference whether player i reports a deviation or not due to the target-payoff modification (cf. Section 5.3, Case 2).

Case B.2. Assume that player i considers a deviation in a punishment master-block $P[j]$, following master-block M_k and initiated by a reported deviation in D_{k-1}.

Case B.2.1: Assume that the deviation is planned for block $R^*[j]$. If player i chooses a different probability distribution over the prescribed actions, it would not be profitable by the target-payoff modification (even if player i is the reporting player, that is, if $i = j$). Obviously, buying information at $R^*[j]$ is useless.

Case B.2.2: Assume that the deviation is planned for block C^*. The proof is basically identical to a regular block C_k.

Case B.2.3: Assume that the deviation is planned for punishment-block P^*. It is clear that buying information is useless. If $i = j^*$ (the player who is being punished) then such deviation cannot be profitable, since player i is supposed to play a best reply against the opponents' mixed actions. If i is one of the punishers, then the target-payoff modification makes him indifferent between any action he could choose. Hence, no deviation can be profitable.

Case B.2.4: Assume that the deviation is planned for block C^{**}. The proof is basically identical to a regular block C_k.

Case C. Assume that player i considers a deviation in Case 2.2.1. So, player i considers a deviation in block C_k of a regular master-block M_k, or in block C^*, P^* or C^{**} of a punishment master-block $P[j]$. By construction, all stages of these blocks are bought by all players, and taken into account at the target-payoff modification, and deviations are punished.

Case D. Assume that player i considers a deviation in Case 2.2.2. So, player i considers a deviation in block $R^*[j]$ in a punishment master-block $P[j]$. By construction, all stages in $R^*[j]$ are bought later at C^*. If $i \neq j$, then by the target-payoff modification, the actions chosen by player i have no influence on his expected discounted payoff. If $i = j$, then punishments make sure that it is optimal for player i to report an opponent (if this is still possible). By the target-payoff modification, player i will be indifferent between any opponent he can name. Hence, no deviation is profitable.

Case E. Assume that player i considers a deviation in Case 2.2.3. So, player i considers a deviation in block D_k of a regular master-block. If $i = j'$ (recall that $P[j']$ was the last punishment-master-block; if no punishment has yet

---

[23] Player i does not know in advance which opponent would detect his deviation. Therefore, with a positive probability, it would be an opponent different from j', which would result in a punishment.

taken place, then $j' = 1$) then he cannot become the player who names the deviator. Hence, due to the target-payoff modification, he has no reason to deviate from the prescribed behavior in Case 2.2.3. We may therefore assume from now on that $i \neq j'$.

Let $\tilde{I}$ be the set of opponents of player i, except j', who, according to player i, should report a deviation in D_k. We distinguish the following cases:

Case E.1. Assume that $\tilde{I}$ is empty. By reporting a deviation, player i would believe that he will be the only reporting player except player j' possibly, and will therefore believe that he will be compensated by the target-payoff modification (see Section 5.3, Case 1). This would give him the same expected discounted payoff as he would get by not reporting. Player i is therefore indifferent between reporting and not reporting a deviation.

Case E.2. Assume that $\tilde{I}$ contains only one opponent, say j. Then, i believes that a deviation has been detected, and believes that j knows that i is aware of this deviation. Consequently, player i is supposed to report a deviation according to the strategy profile. We distinguish two cases:

Case E.2.1. Assume that i believes that his deviation has been detected by a player other than j'. Then, j must be the player who detected the deviation by player i. So, player j would name, with positive probability, player i as the deviator. Hence, player i's only chance to escape from punishment is to report a deviation himself (and name another player as the deviator).

Case E.2.2. Assume that i believes that his deviation has not been detected by a player other than j'. If player i deviates by not reporting at D_k, then he believes that j will be the only player other than j' who reports a deviation in D_k, and that he will name, with a positive probability, player i as the deviator in $R^*[j]$. Hence, player i will be punished, with positive probability, by not reporting this deviation.

Case E.3. Assume that $\tilde{I}$ contains more than one opponent. Then, it does not make a difference whether player i reports a deviation or not due to the target-payoff modification (cf. Section 5.3, Case 2).

<u>*Case F.*</u> Assume that player i considers a deviation in Case 2.2.4. So, player i considers a deviation at U_k or I_k in a regular master-block M_k. If player i would play according to τ_i from C_k on, then, by construction, no deviation can be profitable. If player i would not play according to τ_i from C_k on, such deviation would fall under a previous case.

If there would be at least four players, then an additional idea would be the following: If a deviation is reported, then *all* players who reported a deviation are asked to report a deviator, and the player who has been named at least twice will be punished. The difference is that with four players or more, according to the plan of action, there will always be at least three players who observe a deviation. Hence, even if one of the reporting players would deviate by reporting differently, there would still be at least two players who report correctly.

6. Extensions and concluding remarks

6.1. At least three players, at least two players with at least four actions, other players with at least three actions

In this case, our construction can easily be generalized as follows: In U_k, the players follow the prescribed actions, except for the switching stages at which they randomize equally between two non-prescribed actions. In I_k, C_k and D_k, the two players with at least four actions play according to Table 1. The players with less than four actions use only one action for each of the two possible reports. Since with this construction every player is uncertain at I_k, C_k and D_k about the actions of at least one opponent, every player is forced to buy information as in Table 1. Our main theorem would therefore still hold in this case.

6.2. Three players, three actions

If we have three players, the extension of our result to three actions is problematic for the following reasons: First of all, the construction as summarized in Table 1 cannot be generalized for three actions. The main problem lies in control-block C_k, where players have to send a report based on the observed actions. For four actions it was possible to construct two different probabilistic reports, each one randomizing over two actions, with disjoint supports. It is essential that both reports randomize over at least two actions since otherwise players will not have an incentive to buy this stage. We also need that the two supports are disjoint. For if there would be an action a contained in both supports, then the player could always play a without purchasing information about past stages while being sure that

this will not be detected by his opponents. Clearly, having two disjoint supports of two actions is not possible with only three actions.

For three actions, say a, b and c, one could attempt to overcome this problem by constructing three, instead of two, different codes for reporting at C_k. A first code with support $\{a, b\}$, a second with support $\{b, c\}$ and a third with support $\{a, c\}$. In this way, each code still uses a randomization, and there is no action contained in all three supports. The report to be sent by player i in C_k would still depend on the actions observed by purchasing information about block I_k, and the switching stages t_{jk} for all players j. However, if there is a switching stage t_{jk} which is only chosen by one player $j \neq i$, then player i need not purchase t_{jk} in C_k. The reason is as follows: Since player j has only two possible switching actions, player i can reduce, without buying t_{jk}, the possible reports to two. As these two reports have a common action, player i could safely play this action without buying information about t_{jk}. So, this would not work.

The problem above occurred since players only had two possible switching actions. One could attempt to overcome this problem by letting the players randomize equally over all three actions at the switching stages. By this construction, player i would have an incentive to buy all switching stages, since he can no longer reduce the possible reports to two without buying all opponents' switching stages. However, with probability $1/3$ the switching action of player i coincides with the action prescribed by the target-payoff-vector at that stage. This leads to the following problem in block I_k: Suppose player i has chosen a switching action at stage t_{ik} different from the action prescribed by the target-payoff-vector, and by switching he received a higher payoff. Then, player i could profitably deviate in I_k in the following way: Player i could buy information about the first stage in I_k at the beginning of the second stage in I_k. With a positive probability, each opponent of player i chose a code different from player i's. Hence, with a positive probability, player i may conclude that both opponents have chosen a switching stage different from t_{ik}. In this case, player i may safely report a switching stage $\tilde{t}_{ik} \neq t_{ik}$ in I_k, since t_{ik} will not be bought by the other players and the prescribed action at $\tilde{t}_{ik}$ is also a possible switching action. By doing so, the gain in payoff at t_{ik} would not be compensated by the new target-payoff-vector. If the cost c of buying the first stage in I_k is small enough compared to the gain at stage t_{ik}, then such deviation would be profitable.

6.3. At least four players, three actions

For at least four players and three actions, say a, b and c, the construction in our main theorem could be adapted as follows: Block U_k remains essentially the same, by letting every player choose a switching stage according to the same probability distribution as in Table 1, at which he randomizes between the two non-prescribed actions. In I_k the situation changes a bit, since digit 1 now corresponds to action a and digit 0 corresponds to action b. Hence, the digits no longer correspond to randomizations over two actions, as before. This is necessary since we only have three actions, and hence we cannot construct two randomizations with disjoint supports. In fact, we could also have used this approach for our construction in the main body of the paper. In C_k, every player i still buys the same information as in Table 1, except for the fact that he no longer buys D_{k-1} (as this block will disappear in the new construction), and player i no longer buys the switching stage of player $i + 1$ (if player i is the last player, he does not buy player 1's switching stage). Subsequently, player i either plays a with probability 1 or b with probability 1, based on the observed actions at the switching stages t_{jk} ($j \neq i + 1$) and block C_{k-1}. Block D_k no longer exists. If a player detects a deviation, then he will buy all previous stages, finds out the first deviator, and will always play a non-prescribed action at all further stages, until he notices that all opponents (expect the first deviator) also do so. From that moment on, the first deviator will be punished.

We need four players here in order to guarantee that a deviation at a non-switching stage in U_k will be detected with positive probability. Consider, namely, player i who deviates at a non-switching stage. Then, with positive probability this stage will coincide with player $(i - 1)$'s switching stage, which will be bought by player $i + 1$ (since we have at least four players). Therefore, i's deviation will be detected with positive probability by player $i + 1$.

Notice that by this construction no player is certain about the report to be sent by the opponents in C_k, since he does not buy all switching stages. Therefore, every player has an incentive to buy C_k at C_{k+1}, even though no randomization takes place in C_k. Moreover, every player i should report correctly at C_k since any player $j \neq i$ will know after C_{k+1} which report should have been sent by player i at C_k. Namely, player j has observed all switching stages except the one of player $j + 1$. When at C_{k+1} player j observes the report sent by players $j' \notin \{i, j\}$ at C_k, then

he can deduce the actions played at the switching stage of player $j + 1$. Therefore, player j will know the report that should have been sent by player i at C_k.

6.4. Concluding remarks

1. We wish to mention that the case of two players and two actions for each player is even more complicated. It is unclear, for instance, whether in the two-player prisoners' dilemma one can obtain equilibrium payoffs close to the cooperative outcome. We have the feeling that this is not possible. In any case, one would need a drastically different construction than the one we presented, even if it were possible.

2. In our model, it is crucial that players can buy information about all past stages, and that players always receive information about the actions of all opponents when purchasing a stage. The analysis would be significantly different if players could only buy information about the present stage, or could request information from subgroups of opponents.

Acknowledgments

References

Aumann, R.J., Shapley, L.S., 1994. Long-term competition—A game-theoretic analysis. In: Megiddo, N. (Ed.), Essays in Game Theory. New York, Springer-Verlag, pp. 1–15.

Ben-Porath, E., Kahneman, M., 2003. Communication in repeated games with costly monitoring. Games Econ. Behav. 44, 227–250.

Fudenberg, D., Maskin, E., 1986. Folk theorems for repeated games with discounting and incomplete information. Econometrica 54, 533–554.

Green, E., Porter, R., 1984. Noncooperative collusion under imperfect price information. Econometrica 52, 87–100.

Kandori, M., Obara I., 2004. Endogenous monitoring. Unpublished manuscript, University of Tokyo and UCLA.

Miyagawa, E., Miyahara, Y., Sekiguchi, T., 2004. The Folk theorem for repeated games with observation costs. Discussion Paper Series 23, Kobe University.

Radner, R., 1986. Repeated partnership games with imperfect monitoring and no discounting. Rev. Econ. Stud. 53, 43–58.

Radner, R., Myerson, R., Maskin, E., 1986. An example of a repeated partnership game with discounting and with uniformly inefficient equilibria. Rev. Econ. Stud. 53, 59–70.