Theoretical Economics 13 (2018), 87–113 1555-7561/20180087

High frequency repeated games with costly monitoring

Ehud Lehrer
School of Mathematical Sciences, Tel Aviv University and INSEAD

Eilon Solan
School of Mathematical Sciences, Tel Aviv University

We study two-player discounted repeated games in which one player cannot monitor the other unless he pays a fixed amount. It is well known that in such a model the folk theorem holds when the monitoring cost is on the order of magnitude of the stage payoff. We analyze high frequency games in which the monitoring cost is small but still significantly higher than the stage payoff. We characterize the limit set of public perfect equilibrium payoffs as the monitoring cost tends to 0. It turns out that this set is typically a strict subset of the set of feasible and individually rational payoffs. In particular, there might be efficient and individually rational payoffs that cannot be sustained in equilibrium. We also make an interesting connection between games with costly monitoring and games played between long-lived and short-lived players. Finally, we show that the limit set of public perfect equilibrium payoffs coincides with the limit set of Nash equilibrium payoffs. This implies that our characterization applies also to sequential equilibria.

Keywords. High frequency repeated games, costly monitoring, Nash equilibrium, public perfect equilibrium, no folk theorem, characterization.

JEL classification. C72, C73.

1. Introduction

A key result in the theory of discounted repeated games with full monitoring is the folk theorem, which states that as the discount factor goes to 1, the set of subgame-perfect equilibrium payoffs converges to the set of feasible and individually rational payoffs.[1] The folk theorem implies in particular that a long-term interaction enables efficiency: the efficient and individually rational feasible payoffs can be sustained in equilibrium. This observation is valid when players fully monitor each other's moves

Ehud Lehrer: lehrer@post.tau.ac.il
Eilon Solan: eilons@post.tau.ac.il

This research was supported in part by the Google Inter-University Center for Electronic Markets and Auctions. Lehrer acknowledges the support of the Israel Science Foundation, Grant 963/15. Solan acknowledges the support of the Israel Science Foundation, Grants 212/09 and 323/13. We thank Tristan Tomala, Jérôme Renault, and Phil Reny for useful comments on previous versions of the paper. We are particularly grateful to Johannes Hörner for drawing our attention to the paper by Fudenberg et al. (1990).

[1]To be precise, an additional full-dimensionality condition is required.

https://doi.org/10.3982/TE2627

and, consequently, can enforce any pattern of behavior that results in an individually rational payoff.

In practice, more often than not, players do not perfectly monitor each other's actions, but do obtain signals that depend on their actions. This paper discusses situations where players cannot freely observe the actions taken by their opponents. Rather, players observe other players' actions only if they pay a monitoring fee. For instance, violations of various treaties, such as the Treaty on the Non-Proliferation of Nuclear Weapons or the Convention for the Protection of Human Rights and Fundamental Freedoms, are often difficult to identify. As a result, the international community conducts periodic inspections to ensure that these treaties are kept. Similarly, espionage among countries and industrial espionage with the goal of revealing actions and intent of opponents is a common practice.

Repeated games with costly monitoring have been previously studied in the literature. In such models where the monitoring cost is fixed, namely, independent of the discount factor, the folk theorem has been obtained (see an elaboration below). Our goal in the current paper is to evaluate the robustness of this result to changing monitoring costs. We are particularly interested in the role played by the order of magnitude of monitoring costs relative to stage payoffs. To this end we study a high frequency discrete-time repeated game, a model that can be thought of as a good approximation to a continuous-time game. We consider the impact of small period length in a two-player repeated game, where the inspection cost is small, but of higher order than per period payoffs. This model is relevant when an inspection requires a fixed amount of effort that does not depend on the interaction frequency. It may occur, for instance, when preparing to launch an inspection team or when the inspection itself is expected to take a fixed amount of time and effort that are not affected by the length of the period inspected. An inspection by the tax authority, for example, is highly time consuming, even when it aims to inspect just one single taxpayer in one single year. Another example is a country that uses a collaborator to obtain important information from an enemy country. The employment of a collaborator puts him or her in danger of being captured, thereby inflicting, regardless of the significance of the information obtained, a huge cost on the spying country.

In the presence of inspection costs, it is too costly to monitor the other player at every stage. There are simple equilibria that do not require any inspection, such as playing constantly an equilibrium of the one-shot game. Furthermore, by playing different constant one-stage equilibria in different stages, one can obtain, as a limit of public perfect equilibrium (PPE) payoffs of the repeated game, any point in the convex hull of the one-shot game equilibrium payoffs. A natural question arises as to whether additional (and maybe more efficient) payoffs can be supported by equilibria.

The main objective of the paper is to characterize the limit set of PPE payoffs as the players become more patient. We show that the equilibrium payoffs of a player cannot exceed a certain upper bound determined by the structure of the one-shot game. This implies that costly monitoring typically impedes cooperation: not all the efficient and individually rational payoffs can be sustained by an equilibrium, and so there is an anti-folk theorem.

The goal of anti-folk theorems, such as the one presented here, is to identify the assumptions needed to get a folk theorem. An important insight from our result is that whether the folk theorem applies depends on the magnitude of the monitoring cost relative to the stage-game payoffs. When the monitoring cost is of the same magnitude as the stage-game payoffs, a folk theorem is obtained, while when the monitoring cost is much higher, efficiency is lost and an anti-folk theorem is obtained.

To explain why efficiency is lost, suppose that at some stage of an equilibrium, player i does not play a best response to player j's mixed action. In this case, so as to deter player i from gaining by a deviation that would go unnoticed, player j must monitor player i with a sufficiently high probability. Since monitoring is costly, player j should be later compensated for monitoring player i. Moreover, since the monitoring cost is higher than the contribution of a single stage payoff on the total payoff, when player j monitors player i, her continuation payoff that follows the monitoring must be higher than her expected payoff prior to monitoring.

Now consider player 1's maximal equilibrium payoff in a repeated game with costly monitoring and an equilibrium that supports it. If player 1 monitors player 2 with positive probability at the first stage, his continuation payoff following the monitoring should be higher than his expected payoff prior to the monitoring. In other words, the continuation payoff should be higher than the maximal equilibrium payoff, which is impossible. Consequently, at the first stage, player 1 does not monitor his opponent. This implies that in the first stage, player 2 should not have an incentive to deviate: she already plays a one-shot best response and there is no need for player 1 to inspect her. This reasoning not only shows the connection between player 1's maximal equilibrium payoff and action pairs in which player 2 plays a one-shot best response payoff, it also imposes an upper bound on player 1's equilibrium payoffs and thereby restricts efficiency. Player 2's equilibrium payoffs are also subject to a similar upper bound.

It turns out that the upper bound over equilibrium payoffs thus obtained is similar to the upper bound over equilibrium payoffs in case of long-lived players playing against a sequence of short-lived players (see Proposition 3 in Fudenberg et al. (1990), which is analogous to our Theorem 1). In such an interaction, a short-lived player has only short-term objectives, threats of punishment are not effective against such a player, and he therefore always plays in equilibrium a one-shot game best response. Consequently, the maximal equilibrium payoff of a long-lived player is characterized by action profiles where the short-lived players play a best response. This observation comprises an interesting similarity with our model: the maximal equilibrium payoff of each player in our model is precisely the bound on the long-lived player defined in Fudenberg et al. (1990).

Despite the similarities, there are two essential differences between the results in the two models. First, the restricted inefficiency in a game with short-lived players is a consequence of the fact that these players consider only the immediate stage interaction. In our model, in contrast, both players have long-run objectives. It is only when the expected payoff of one player from this stage and on is equal to his maximal equilibrium payoff that the other player behaves like a short-lived player. Moreover, even in this case, this behavior is temporary and applies only to the current stage of the game. Once the expected payoff of the player from this stage and on falls below his maximal equilibrium payoff, the opponent's behavior is no longer the behavior of a short-lived player. The second difference is that in Fudenberg et al. (1990), the upper bound on payoffs applies only to the long-lived player, while in our model it further restricts efficiency since it applies to both players.

In constructing equilibria, monitoring is crucial both to sustain and to enforce equilibrium payoffs. Specifically, monitoring serves three different purposes.

1. Monitoring the other player with sufficiently high probability, coupled with a threat of punishment, ensures that in equilibrium the other player will not deviate to an action that he is not supposed to play.

2. When a player plays a mixed action, different actions played with positive probability may yield different payoffs. To make the player indifferent as to which action he takes, different continuation payoffs must be attached to different actions. Monitoring is used to enable the players to coordinate the continuation payoffs. When a player is supposed to play a mixed action, he is monitored with a positive probability, and in case he is monitored, the continuation payoff is set in such a way as to make the player indifferent between his actions.

3. Since monitoring is costly, in equilibrium a player can monitor the other player so as to burn money. He will do so because otherwise he will be punished, and the resulting payoff would be worse. This possibility of forced monitoring enables one to design relatively low continuation payoff. For instance, suppose that a player prescribed to play a certain mixed action is monitored, and it turns out that the realized pure action yields him a high payoff. This player can be instructed later on to monitor and pay the monitoring cost, and thereby reduce his own payoff.

In our model, monitoring is common knowledge. In particular, both players know its outcome. This implies that the problem of characterizing the set of equilibrium payoffs is recursive. Indeed, our proof method is recursive in nature: we have a "conjecture" about the limit set of PPE payoffs, and for each point in this set, we provide a proper one-shot game and continuation payoffs in the set that render it an equilibrium.

The literature on games with imperfect monitoring

When the magnitude of monitoring costs equals that of stage payoffs, a repeated game with costly monitoring can be recast as a game with imperfect monitoring, which is surveyed in, e.g., Pearce (1992), Mailath and Samuelson (2006), and Mertens et al. (2016). Indeed, the choice weighed by each player at every stage is composed of two components: (a) which action to play and (b) whether to monitor the other player. The payoff function can be adapted accordingly: in case no monitoring is performed, the stage payoff coincides with the original payoff and is equal to the original payoff minus the monitoring cost otherwise. In our setup, the monitoring cost depends on the discount factor and is significantly larger than the stage payoff, and therefore the game cannot be modelled as a repeated interaction with imperfect monitoring: monitoring cannot be considered as a regular action of an extended base game.

Undiscounted repeated games with imperfect monitoring have been studied by Lehrer (1989, 1990, 1991, 1992). Abreu et al. (1990) analyzed discounted games and used dynamic programming techniques to characterize the set of public equilibrium payoffs. Fudenberg et al. (1994) provided conditions that guarantee that any feasible and individually rational payoff is a perfect equilibrium payoff when players are sufficiently patient. Fudenberg and Levine (1994) characterized the limit set of public perfect equilibrium payoffs in the presence of both public and private signals as the discount factor goes to 1. Compte (1998) and Kandori and Matsushima (1998) proved a folk theorem for repeated games with communication and independent private signals. Hörner et al. (2011) extended the characterization to stochastic games.

Several authors studied specifically repeated games with costly observations. Ben-Porath and Kahneman (2003) studied a model in which, at the end of every stage, each player can pay a fixed amount and observe the actions just played by a subset of other players. They proved that if the players can communicate, the limit set of sequential equilibrium payoffs when players become patient is the set of feasible and individually rational payoffs. Miyagawa et al. (2008) assumed that monitoring decisions are not observed by others, that players have a public randomization device, and that they observe a stochastic signal that depends on other players' actions even if they do not purchase information. They proved that under full dimensionality condition, the folk theorem is still obtained. In the model studied by Flesch and Perea (2009), players can purchase information on actions played in past stages and in the current stage. They proved that in case at least three players (resp. four players) are involved and each player has at least four actions (resp. three actions), a folk theorem for sequential equilibria holds.

The results attained by the last three papers mentioned is different from ours. They obtain the standard folk theorem, while we do not. The reason for this difference is that in their models, the monitoring cost is bounded. As mentioned above, this kind of model is a special case of repeated games with imperfect monitoring.

Another related paper in a different strand of literature is Lipman and Wang (2009), who studied repeated games with switching costs. In this model a player has to pay a fixed cost whenever playing different actions in two consecutive stages. Similarly to our cost structure, the switching cost in Lipman and Wang (2009) is much higher than the stage payoff. Nevertheless they obtain a folk theorem.

The structure of the paper

The model is presented in Section 2. Section 3 provides the upper bounds on the payers' payoffs and a no-folk-theorem result. Section 4 characterizes the set of public perfect equilibrium payoffs, while Section 5 provides the main ideas of the equilibrium construction. Final comments are given in Section 6. The proofs appear in the Appendix, available in a supplementary file on the journal website, http://econtheory.org/supp/2627/supplement.pdf.

2. The model

2.1 *The base game*

Let $G = (\{1, 2\}, A_1, A_2, u_1, u_2)$ be a two-player one-shot *base game* in strategic form. The set of players is $\{1, 2\}$, A_i is the finite set of player i's actions, and $u_i : A \to \mathbb{R}$ is his payoff function, where $A := A_1 \times A_2$. As usual, the multi-linear extension of u_i is still denoted by u_i. For notational convenience, we let j denote the player who is not i.

The minmax value (in mixed strategies) of player i in the base game is given by[2]

$$v_i := \min_{\alpha_j \in \Delta(A_j)} \max_{\alpha_i \in \Delta(A_i)} u_i(\alpha_i, \alpha_j).$$

We assume without loss of generality that the maximal payoff in absolute values, $\max_{i=1,2} \max_{a \in A} |u_i(a)|$, does not exceed 1. Denote the *minmax point* by $v := (v_1, v_2)$. A payoff vector $x \in \mathbb{R}^2$ is *individually rational* (resp. *strictly individually rational*) for player i if $x_i \geq v_i$ (resp. $x_i > v_i$). Denote by F the set of all vectors in $\mathbb{R}^2$ dominated by a feasible vector in the base game:[3]

$$F := \{x \in \mathbb{R}^2 : \exists y \in \operatorname{conv}\{u(a), a \in A\} \text{ such that } y \geq x\}.$$

Since monitoring is costly, players can use the monitoring option to burn money. Therefore, the set of feasible payoff vectors in the repeated game is the set of vectors dominated by feasible payoffs in the base game.

2.2 *The repeated game*

We study a repeated game in discrete time, denoted $G(r, c, \Delta)$, which depends on three parameters, $r \in (0, 1)$, $c > 0$, and $\Delta > 0$, and on the base game G. This game is described as follows.

1. The base game G is played over and over again.
2. The duration between two consecutive stages is Δ.
3. The discount factor is r.
4. At every stage of the game each player chooses an action in the base game and whether to monitor the action chosen by the other player. Monitoring the other player's action costs c and becomes common knowledge. We denote by O_i (resp. NO_i) the choice of player i to monitor (or observe; resp. not to observe) player j's action.

A *private history of player i* at stage n ($n \in \mathbb{N}$) consists of (a) the sequence of actions he played in stages $1, 2, \ldots, n-1$, (b) the stages in which player j monitored him, (c) the stages in which he monitored player j, and (d) the actions that player j played in those

[2] For every finite set X, we denote by $\Delta(X)$ the set of probability distributions over X.

[3] Let $x, y \in \mathbb{R}^2$. We denote $y \geq x$ if $y_i \geq x_i$ for each $i = 1, 2$. In this case we say that x is *dominated* by y. The vector x is *strictly dominated* by y when $y_i > x_i$ for each $i = 1, 2$.

stages. Denote by $H_i(n-1)$ the set of all such private histories. The set $H_i(n-1)$ consists of all player i's information sets before making a decision at stage n. Note that $H_i(n-1)$ is a finite set. A *public history* at stage n consists of (a) the stages in which each player monitored the other player prior to stage n and (b) the actions that the monitored player took in these stages. The public history is commonly known to both players. Denote by $H^P(n-1)$ the set of public histories at stage n. Let $\mathcal{F}^{n-1}$ be the σ-algebra defined on the space of infinite plays H^∞ and spanned by the set of all public histories of length $n-1$.

A *pure (resp., public pure) strategy* of player i is a function that assigns two components to every private (resp., public) history in $H_i(n-1)$ (resp., $H^P(n-1)$): an action in A_i to play at stage n and a binary variable, either O_i or NO_i, that indicates whether player i monitors player j at stage n.

A *behavior* (resp. public behavior) strategy of player i is a function that assigns a probability distribution over $A_i \times \{O_i, NO_i\}$ for every stage n and every private (resp. public) history in $H_i(n-1)$ (resp. $H^P(n-1)$). In our construction we only use public behavior strategies in which these distributions are product distributions. That is, the action played at stage n is conditionally independent of the decision whether to monitor at that stage. Since the players have perfect recall, by Kuhn's theorem every public behavior strategy is strategically equivalent to a mixed public strategy and vice versa.

Every pair of strategies (σ_1, σ_2) induces a probability distribution $\mathbb{P}_{\sigma_1,\sigma_2}$ over the set of infinite plays H^∞, supplemented with the σ-algebra generated by all finite cylinders. We denote by $\mathbb{E}_{\sigma_1,\sigma_2}$ the corresponding expectation operator. Denote by α_i^n player i's mixed action at stage n, and $\alpha^n = (\alpha_1^n, \alpha_2^n)$. The total (expected) payoff to player i when the players use the strategy pair (σ_1, σ_2) is

$$U_i(\sigma_1, \sigma_2) := \mathbb{E}_{\sigma_1,\sigma_2}\left[(1-r^\Delta)\left(\sum_{n=1}^{\infty} r^{\Delta(n-1)} u_i(\alpha^n)\right) - c\left(\sum_k r^{\Delta(\tau_i^k - 1)}\right)\right], \tag{1}$$

where $(\tau_i^k)_{k\in\mathbf{N}}$ are the stages in which player i monitors player j.

It is worth noting that the contribution of the stage payoff to the total discounted payoff depends on the duration Δ between stages, and is equal to $(1-r^\Delta)u_i(\alpha^n)$. The discounted value of the nth stage payoff is therefore equal to $(1-r^\Delta)r^{\Delta(n-1)}u_i(\alpha^n)$. Conversely, the monitoring cost is much higher than the stage payoff. It is constant and does not depend on the duration between stages. This is why the cost of the kth observation, which is performed at stage τ_i^k, is multiplied by $r^{\Delta(\tau_i^k-1)}$ and not by $(1-r^\Delta)$. The difference between the nature of the stage payoff and that of the monitoring cost is the point where our model departs from the literature.

2.3 *Equilibrium*

A pair of strategies is a *(Nash) equilibrium* if no player can increase his total payoff by deviating to another strategy. A *public equilibrium* is an equilibrium in public strategies. In such an equilibrium, no player can profit by deviating to any strategy, public or not public. A *public perfect equilibrium* is a pair of public strategies that induces an equilibrium in the continuation game that starts after any public history. Let $\mathrm{NE}(r, c, \Delta)$ be the set of Nash equilibrium payoffs in the game $G(r, c, \Delta)$ and let $\mathrm{PPE}(r, c, \Delta)$ be the set of public perfect equilibrium payoffs of this game.

Define

$$\mathrm{NE}^*(r) = \limsup_{c\to 0} \limsup_{\Delta\to 0} \mathrm{NE}(r, c, \Delta),$$
$$\mathrm{PPE}^*(r) = \limsup_{c\to 0} \limsup_{\Delta\to 0} \mathrm{PPE}(r, c, \Delta).$$

These are the limit sets of Nash equilibrium payoffs and public perfect equilibrium payoffs as both the duration between stages and the observation cost go to 0, and the former goes to 0 faster than the latter.

By definition, $\mathrm{PPE}(r, c, \Delta) \subseteq \mathrm{NE}(r, c, \Delta)$ for every discount factor r, every observation cost c, and every duration Δ, and therefore $\mathrm{PPE}^*(r) \subseteq \mathrm{NE}^*(r)$. Our main result characterizes these sets in terms of the base game. It turns out that under a weak technical condition these two sets coincide.

Playing a Nash equilibrium of the base game at every stage and after every history, without monitoring each other, is a stationary equilibrium of the game $G(r, c, \Delta)$. We therefore conclude that the set $\mathrm{PPE}(r, c, \Delta)$ contains the set NE of Nash equilibrium payoffs of the base game. By partitioning the set of stages into disjoint subsets, and playing the same Nash equilibrium in all stages of the subset, without monitoring the other player, we construct an equilibrium payoff in the convex hull of NE. When $r^\Delta > \frac{1}{2}$, this construction can yield any vector in the convex hull of NE. We thus obtain the following lemma.

Lemma 1. *For every $r > 0$, the set $\mathrm{PPE}^*(r)$ contains the convex hull of the set* NE.

A *maxmin mixed action* of player i in the base game is any mixed action $\alpha_i \in \Delta(A_i)$ that satisfies $u_i(\alpha_i, a_j) \geq v_i$ for every $a_j \in A_j$. By repeating his maxmin mixed action in the base game and not monitoring the other player, player i guarantees a payoff v_i in the repeated game $G(r, c, \Delta)$. We conclude with the following result.

Lemma 2. *For every $r, c, \Delta > 0$ and $x \in \mathrm{NE}(r, c, \Delta)$, one has $x \geq v$.*

3. No folk theorem

In this section, we show that the folk theorem does not hold in games with costly monitoring. In Section 3.1, we present two quantities M_1 and M_2, and in Section 3.2, we show that M_i is an upper bound to player i's payoff in $\mathrm{NE}(r, c, \Delta)$. In some games these bounds are restrictive, in the sense that they are lower than the highest payoff of player i in the set of feasible and individually rational payoffs. In particular, it may happen that the set $\mathrm{NE}(r, c, \Delta)$ may be disjoint from the Pareto frontier of $F \cap V$. In subsequent sections, we characterize the sets $\mathrm{NE}^*(r)$ and $\mathrm{PPE}^*(r)$ using the quantities M_1 and M_2.

		Player 2	
		D	C
Player 1	D	1, 1	4, 0
	C	0, 4	3, 3

FIGURE 1. The prisoner's dilemma.

3.1 *Best response and the index M_i*

We say that player i *plays a best response* at the mixed-action pair $\alpha = (\alpha_1, \alpha_2)$ if

$$u_i(\alpha_1, \alpha_2) = \max_{a_i \in A_i} u_i(a_i, \alpha_j).$$

Player i is *indifferent* at $\alpha = (\alpha_1, \alpha_2)$ if for every action a_i such that $\alpha_i(a_i) > 0$,

$$u_i(\alpha_i, \alpha_j) = u_i(a_i, \alpha_j).$$

We now define two indices M_1 and M_2 that play a major role in our characterization. Let

$$M_i := \max\Big\{\min_{a_i:\, \alpha_i(a_i)>0} u_i(a_i, \alpha_j) \colon (\alpha_1, \alpha_2) \in \Delta(A_1) \times \Delta(A_2) \text{ and } \alpha_j \text{ is a best response to } \alpha_i\Big\}. \tag{2}$$

To explain the definition in (2), consider M_2. Let α_2 be a mixed action of player 2 and let α_1 be a best response of player 1 to α_2. By playing α_2, player 2 does not necessarily optimize against α_1, implying that any pure action in the support of α_2 might induce a different payoff for player 2. We focus on the minimum among these payoffs, which is a function of the pair (α_1, α_2). The index M_2 is the maximum of all these minimal numbers, over all pairs (α_1, α_2), where α_1 is a best response of player 1 to α_2.

The next example illustrates the quantity M_2 in the prisoner's dilemma.

EXAMPLE 1 (The prisoner's dilemma). The prisoner's dilemma is given by the base game that appears in Figure 1.

We calculate M_2 for this game. Fix a mixed action α_2 of player 2. The best response of player 1 to α_2 is D and $\min_{a_2 \in \text{supp}(\alpha_2)} u_2(D, a_2)$ is either 1 or 0, where $\text{supp}(\alpha_i) := \{a_i \in A_i \colon \alpha_i(a_i) > 0\}$. The maximum over these minima is 1, implying that $M_2 = 1$. ◊

By the definition of M_i, if α_j is a best response to α_i, then

$$M_i \geq u_i(a_i, \alpha_j) \quad \text{for at least one action } a_i \in \text{supp}(\alpha_i).$$

Consequently, M_i is at least as high as the payoff of player i in any equilibrium of the base game. Formally, for every Nash equilibrium α of the base game,

$$M_i \geq u_i(\alpha), \quad i \in \{1, 2\}.$$

		Player 2 L	C	R
Player 1	T	1, 1	3, 0	0, 0
	I	0, 0	2, 2	0, 0
	B	0, 3	0, 0	2, 2

FIGURE 2. The game in Example 2.

The following example shows that M_i might be strictly higher than player i's payoffs in all Nash equilibria.

EXAMPLE 2. Consider the 3×3 base game that appears in Figure 2.

By an iterative elimination of pure strategies, one deduces that the action pair (T, L) is the unique equilibrium of the base game, and it yields the payoff $(1, 1)$. Since C is a best response to I, we deduce that $M_1 \geq 2$, and since B is a best response to R, we deduce that $M_2 \geq 2$. ◊

The significance of M_i becomes apparent in Theorem 1 below. It states that M_i is the upper bound of player i's payoffs in the repeated game $G(r, c, \Delta)$. The intuition is as follows. We first explain why the definition of M_i concerns mixed actions at which player j plays a best response. When player i monitors player j, the former incurs a significant monitoring cost (recall that c is significantly larger than Δ). Consequently, in equilibrium, player i's continuation payoff after monitoring must be higher than his expected payoff before doing so. This implies that in equilibrium that supports player i's maximal payoff in the set $\mathrm{NE}(r, c, \Delta)$, he does not monitor his opponent at the first stage, because otherwise his continuation payoff following the monitoring would exceed the maximal payoff. Therefore player j must have, in equilibrium, no incentive to deviate in the first stage, meaning that he must play a one-shot best response.

We now explain why M_i is indeed an upper bound of the set of equilibrium payoffs. Assume by contradiction that player i's maximal payoff in the set $\mathrm{NE}(r, c, \Delta)$, denoted x_i^*, is strictly higher than M_i. Denote by $\alpha = (\alpha_1, \alpha_2)$ the mixed action pair that the players play at the first stage of an equilibrium that supports x_i^*. As we saw above, player j plays a best response at α. Consider now the event that player i plays at the first stage a pure action a_i^* that minimizes the stage payoff $u_i(a_i, \alpha_j)$ among the actions a_i such that $\alpha_i(a_i) > 0$. By the definition of M_i we have $u_i(a_i^*, \alpha_j) \leq M_i < x_i^*$. Since a_i^* is played with positive probability at the first stage, x_i^* is a weighted average of $u_i(a_i^*, \alpha_j)$ and the continuation payoff that follows it. This continuation payoff is also an equilibrium payoff and therefore cannot exceed x_1^*. We obtain that x_1^* is a weighted average of two smaller numbers, of which one is strictly smaller. This is a contradiction.

3.2 *Bounding the set of Nash equilibria*

In this subsection, we consider fixed $r, c, \Delta > 0$. Since the set of strategies is compact and the payoff function is continuous over the set of strategy pairs, one obtains the following result.

Lemma 3. *The set* $\mathrm{NE}(r, c, \Delta)$ *of Nash equilibrium payoffs in the repeated game is compact.*

The following theorem states that when Δ is sufficiently small, player i's equilibrium payoff cannot exceed M_i. In particular, it means that not all feasible and individually rational payoffs are equilibrium payoffs (i.e., not all of them are in $\mathrm{NE}(r, c, \Delta)$): costly monitoring typically impairs efficiency.

Theorem 1. *Fix* $\Delta > 0$*,* $i \in \{1, 2\}$*, and* $x \in \mathrm{NE}(r, c, \Delta)$*. If* $\Delta < \frac{\ln(1-\frac{c}{1-x_i})}{\ln(r)}$*), then* $x_i \leq M_i$*.*

Example 1 (Revisited). As mentioned before, in the prisoner's dilemma $M_1 = M_2 = v_1 = v_2 = 1$. Since, by Theorem 1, any Nash equilibrium payoff cannot exceed M_i, we obtain $\mathrm{NE}(r, c, \Delta) = \{(1, 1)\}$ provided that Δ is sufficiently small. In other words, in the repeated prisoner's dilemma, mutual defection is the only equilibrium payoff in the presence of a high monitoring fee. The intuition behind this result is that to implement a payoff that is not $(1, 1)$, at least one player, say player 2, must play the dominated action C. This implies that to deter a deviation to D, player 1 has to monitor player 2 with a positive probability. Whenever player 1 monitors player 2, he should be compensated by a higher continuation payoff for having to bear the monitoring cost. The only circumstance whereby the continuation payoff may compensate player 1 is in case player 2 plays the dominated action C with a higher probability. In that case, player 1 must continue monitoring player 2 with positive probability. It might therefore happen, although with a small probability, that player 1 will have a long stretch of stages in which he monitors player 2, the continuation payoff of player 1 will keep increasing and eventually exceed 4, which is impossible. ◊

Proof of Theorem 1. We prove the theorem for $i = 1$. By Lemma 3, the set $\mathrm{NE}(r, c, \Delta)$ is compact. Let x^* be a payoff vector in $\mathrm{NE}(r, c, \Delta)$ that maximizes player 1's payoff. That is, $x^* \in \mathrm{argmax}\{x_1 : x \in \mathrm{NE}(r, c, \Delta)\}$. Assume to the contrary that $M_1 < x_1^*$. When Δ is sufficiently small, we obtain a contradiction.

Consider an equilibrium σ^* that supports x^*, and denote by $\alpha = (\alpha_1, \alpha_2) \in \Delta(A_1) \times \Delta(A_2)$ the mixed-action pair played under σ^* at the first stage. For every action $a_1 \in A_1$, denote by $I_1(a_1)$ the event that player 1 plays the action a_1 and monitors player 2 at the first stage. Denote by $I_1 := \bigcup_{a_1 \in A_1} I_1(a_1)$ the event that player 1 monitors player 2 at the first stage. Let z_1 be player 1's continuation payoff from stage 2 onward, conditional on his information following stage 1.

The proof is divided into two cases.

Case 1: $\mathbb{P}_{\sigma^*}(I_1) > 0$. Since σ^* is an equilibrium, the expected payoff of player 1 conditional on the event that he monitors player 2 at the first stage must be equal to x_1^*. Furthermore, the event I_1 is common knowledge. If both players monitor each other at the first stage, then the actions of both players are known to both and the continuation play is a Nash equilibrium (of the repeated game). If only player 1 monitors at the first stage, an event that is known to both players, then the expected play following the first stage and conditional on the action of player 2 at the first stage is an equilibrium. Consequently, the expectation of z_1 conditional on I_1 and a_2^1 is at most x_1^*, that is, $\mathbb{E}_{\sigma^*}[z_1|I_1, a_2^1] \leq x_1^*$. We therefore deduce that

$$x_1^* = \mathbb{E}_{\sigma^*}\big[(1-r^\Delta)u_1(\alpha) - c + r^\Delta z_1|I_1\big] \leq 1 - r^\Delta - c + r^\Delta x_1^*.$$

This inequality is violated when $\Delta < \frac{\ln(1-c/(1-x_1^*))}{\ln(r)}$.

Case 2: $\mathbb{P}_{\sigma^*}(I_1) = 0$. Since player 1 does not monitor player 2 at the first stage, α_2 is a best response at α. Otherwise, player 2 would have a profitable deviation at the first stage that would go unnoticed. The definition of M_1 implies that $M_1 \geq \min_{a_1 \in \text{supp}(\alpha_1)} u_1(a_1, \alpha_2)$. Denote by $a_1^* \in \text{supp}(\alpha_1)$ an action that attains the minimum. Since by assumption $M_1 < x_1^*$, one has $u_1(a_1^*, \alpha_2) < x_1^*$.

We claim that $\mathbb{E}_{\sigma^*}[z_1|a_1^*] \leq x_1^*$. If player 2 did not monitor player 1, then each player's play after the first stage is independent of his opponent's action at the first stage, and the expected continuation play is a Nash equilibrium. If player 2 monitors player 1, then as in Case 1 the expected play after the first stage conditioned on the action of player 1 at the first stage is an equilibrium. We thus conclude that $\mathbb{E}_{\sigma^*}[z_1|a_1^*] \leq x_1^*$, and therefore

$$\begin{aligned} x_1^* &= (1-r^\Delta)u_1(a_1^*, \alpha_2) + r^\Delta \mathbb{E}_{\sigma^*}[z_1|a_1^*] \\ &< (1-r^\Delta)x_1^* + r^\Delta x_1^* \leq x_1^*, \end{aligned}$$

a contradiction. □

4. The main result: Characterizing the set of public perfect equilibrium payoffs

The set of individually rational payoff vectors that are (a) dominated by a feasible point and (b) yield to each player i at most M_i, is denoted by

$$F^M := \{x \in F : v_1 \leq x_1 \leq M_1 \text{ and } v_2 \leq x_2 \leq M_2\}.$$

Theorem 1 and Lemma 2 imply that $\text{NE}(r, c, \Delta) \subseteq F^M$, provided that Δ is sufficiently small, and consequently the set $\text{PPE}(r, c, \Delta)$ is a subset of F^M for every Δ sufficiently small. Our main result states that the sets $\text{NE}(r, c, \Delta)$ and $\text{PPE}(r, c, \Delta)$ are close to F^M, provided that c and Δ are sufficiently small. This result implies in particular that the bound M_i, established in Theorem 1, is tight.

We now define the closeness concept between the sets that we use. A set K of payoff vectors is an *asymptotic set of Nash equilibrium payoffs* (resp. of *PPE payoffs*) if any point in the set is close to a point in $\text{NE}(r, c, \Delta)$ (resp. $\text{PPE}(r, c, \Delta)$) for every c and Δ small enough and every discount rate r.

Definition 1. A set $K \subseteq \mathbb{R}^2$ is an *asymptotic set of Nash equilibrium payoffs* if, for every $r > 0$ and every $\epsilon > 0$, there is $c_\epsilon > 0$ such that for every $c \in (0, c_\epsilon]$ there is $\Delta_{c,\epsilon,r} > 0$ such that for every $\Delta \in (0, \Delta_{c,\epsilon,r})$ we have

$$\max_{y \in K} \min_{x \in \text{NE}(r,c,\Delta)} d(x, y) \leq \epsilon.$$

Player 2

		L	R
Player 1	T	$0, 0$	$0, -1$
	B	$-1, 3$	$0, 2$

FIGURE 3. A game where $M_1 = v_1$.

The set K is an *asymptotic set of PPE payoffs* if an analogous condition holds with respect to the set $\mathrm{PPE}(r, c, \Delta)$.

Note that Definition 1 concerns only one direction of the Hausdorff distance: it requires that any point in K is close to a Nash equilibrium payoff (or a PPE payoff), but not vice versa.

THEOREM 2. *If $M_1 > v_1$ and $M_2 > v_2$, then for every discount factor $r \in (0, 1)$ the set F^M is an asymptotic set of Nash equilibrium payoffs and an asymptotic set of PPE payoffs. In particular,* $\mathrm{NE}^*(r) = \mathrm{PPE}^*(r) = F^M$.

The proof of Theorem 2 appears in Section 5 and in the Appendix. As the following example shows, the condition in Theorem 2 requiring that $M_1 > v_1$ and $M_2 > v_2$ cannot be disposed of.

EXAMPLE 3. Consider the 2×2 base game illustrated in Figure 3.

The minmax value of both players is 0. Since the maximum payoff of player 1 is 0, we deduce that $M_1 = 0$, implying that $M_1 = v_1$. Since B is a best response to R and not to L, it follows that $M_2 = 2$. In particular, the set F^M is the interval between $(0, 0)$ and $(0, 2)$. We argue that the unique equilibrium payoff in the repeated game $G(r, c, \Delta)$ is $(0, 0)$, implying that the conclusion of Theorem 2 does not hold. Indeed, since $v_1 = 0$ and the maximal payoff of player 1 is 0, his payoff in every Nash equilibrium of $G(r, c, \Delta)$ as well as his continuation payoff after any public history is 0. The action L strictly dominates the action R, and therefore whenever player 2 plays R with positive probability, he must be monitored by player 1. However, player 1 cannot be compensated for monitoring; hence in equilibrium player 2 always plays L. Since $u_1(B, L) = -1$, player 1 always plays T: in equilibrium the players repeatedly play (T, L) and, consequently, the unique equilibrium payoff is $(0, 0)$, as claimed. ◊

REMARK 1. We assumed that the monitoring fee c is the same for both players. The results remain the same if the monitoring fees of the two players are different, provided that the duration Δ is significantly smaller than both. That is, for every $c_1, c_2 > 0$ sufficiently small, the set of equilibrium payoffs of the two-player repeated game $G(r, c_1, c_2, \Delta)$, in which the monitoring costs of the two players are c_1 and c_2, is close to the set F^M, provided that Δ is sufficiently close to 0. In fact, in our proof it is more convenient to assume that the monitoring fees of the players differ.

The sets of Nash and public perfect equilibrium payoffs in $G(r, c_1, c_2, \Delta)$ are denoted by $\mathrm{NE}(r, c_1, c_2, \Delta)$ and $\mathrm{PPE}(r, c_1, c_2, \Delta)$, respectively.

5. The structure of the equilibrium

Theorem 1 implies that the set $\mathrm{NE}^*(r)$ is included in F^M. To complete the proof of Theorem 2, it remains to prove that $\mathrm{PPE}^*(r)$ contains F^M. We first prove that $\mathrm{NE}^*(r) \supseteq F^M$ by constructing equilibria in which detectable deviations trigger indefinite punishment. We then see how the indefinite punishment can be replaced by a credible threat, implying that $\mathrm{PPE}^*(r) \supseteq F^M$.

At a technical level our proof uses a classical technique. We identify sets X of payoff vectors that have the following property. Every $x \in X$ is an equilibrium of a one-shot game whose payoffs are composed as a payoff from the base game plus a continuation payoff from X itself. We start with a small set X and expand it until we obtain a set close to F^M. The novelty of the proof is in the burning-money process that we proceed now to introduce. This process allows one to decompose the continuation payoff into two parts: a target continuation payoff and the gap between the actual continuation payoff and the target payoff. This gap is precisely the amount the players must burn. While the recursive calculation of continuation payoffs, based on the actual play in the previous stage, is rather complicated, each of the two parts can be easily calculated in a recursive way. The decomposition of the continuation payoff into two parts significantly simplifies the construction of equilibria in the current model, and may be useful in other models as well.

5.1 *Liability and burning-money processes*

To simplify the computations in our construction, we apply a positive affine transformation on the payoffs. Applying such an affine transformation to the player's payoffs in the base game does not change the strategic considerations of the players. However, it does change their monitoring fees and no longer allows us to assume that the monitoring costs are identical for both players. We thus assume from now on that the monitoring costs differ and we denote the monitoring cost of player i by c_i (see Remark 1).

In our construction, players monitor each other for two purposes. First, monitoring is aimed to deter players from deviating. This type of monitoring takes place at random stages. Second, monitoring is used to establish continuation payoffs that ensure that players are indifferent between their prescribed actions. This type of monitoring takes place at known stages.

To implement the second purpose, we introduce burning-money processes. The value of the burning-money process at stage n, which is called the player's *debt*, represents the amount that the player has to burn from stage n onward. This amount is measurable with respect to the public history at that stage, and thus each player knows the other player's debt. Moreover, each player can verify whether the other player burned money as required. The nature of the burning-money process is that as long as the debt is smaller than c_i, the debt is deferred to the next period, and due to discounting, it increases. This happens until the debt exceeds c_i. At this point in time, player i has to monitor player j and as a result, his debt is reduced by c_i. Failing to do so triggers a punishment. The debt might also increase due to other reasons. This might happen when, under equilibrium, a player plays with positive probability two actions that yield different stage payoffs. To ensure that the player is indifferent between the two actions, his debt increases when he is monitored and plays the higher-payoff action.

The definition of the debt process relies on liability processes.

Definition 2. A *liability process* is a nonnegative stochastic process $\xi = (\xi^n)_{n\in\mathbf{N}}$ such that ξ^n is measurable with respect to $\mathcal{F}^{n+1}$ for every $n \in \mathbf{N}$.

The liability is meant to stand for the additional debt that a player incurs at stages in which he is monitored. The role of the liability process is to make all actions played by the monitored player payoff-wise equivalent to him. This is the reason why the liability at stage n depends on the play at that stage and, therefore, ξ_n is $\mathcal{F}^{n+1}$-measurable.

Definition 3. Let $\xi_i = (\xi_i^n)_{n\in\mathbf{N}}$ be a liability process of player i. A *burning-money process based on* ξ_i is a stochastic process $D_i = (D_i^n)_{n\in\mathbf{N}}$ that satisfies the following properties:

- We have $D_i^1 \geq 0$: the initial debt is a nonnegative real number.
- If $D_i^n \geq c_i$, then $D_i^{n+1} = \frac{D_i^n - c_i + \xi_i^n}{r^\Delta}$. The interpretation is that at a stage in which the debt exceeds c_i, player i has to monitor the other player and incurs a cost of c_i, thereby reducing his debt by this amount. The debt D_i^{n+1} is obtained by adding the liability ξ_i^n to the revised debt, and the total is divided by the discount rate r^Δ.
- If $D_i^n < c_i$, then $D_i^{n+1} = \frac{D_i^n + \xi_i^n}{r^\Delta}$. When the debt is below c_i, no mandatory inspection takes place, the liability ξ_i^n is added to the current debt, and the total is divided by the discount rate.

Note that the debt D^n at stage n depends only on the history up to (and including) stage n: D_i^n is measurable with respect to $\mathcal{F}^n$. This implies that at the beginning of stage n the debts of both players are common knowledge. Moreover, the debts are always nonnegative.

In our construction, the liability of player i at stage n is at most $2(1 - r^\Delta)$. Here 2 is the maximal difference between two stage payoffs, and $(1 - r^\Delta)$ is the weight of a single stage payoff. Consequently, player i's debt is bounded by $\frac{c_i + 2(1-r^\Delta)}{r^\Delta}$.

5.2 *Monitoring to detect deviations*

Let $\alpha = (\alpha_1, \alpha_2)$ be a pair of mixed actions played at some stage. When player 1 is indifferent at α and α_1 is not a best response at α, the only way player 1 can gain is by deviating to an action outside the support of α_1. Suppose that player 2 monitors player 1 with probability p. A threat to punish player 1 down to his minmax level in case a deviation is detected is effective if the expected loss due to the punishment is greater than the potential gain: $2(1 - r^\Delta) < p \cdot r^\Delta(x_1 - v_1)$, where x_1 is player 1's expected continuation payoff when he is monitored and no deviation occurs. It follows that to deter deviations to actions outside the support of α_1, we need to set the per-stage probability of monitoring p to satisfy

$$p > \frac{2(1-r^{\Delta})}{r^{\Delta}(x_1 - v_1)}. \tag{3}$$

An analogous inequality holds when player 1 tries to deter deviations of player 2. Note that $\lim_{\Delta \to 0} \frac{1-r^{\Delta}}{\Delta} = -\ln(r)$, and thus the probability that a player is monitored should be larger than $\frac{2\Delta(-\ln(r))}{x_1 - v_1}$, which is of magnitude Δ.

5.3 *The general structure of the equilibrium*

In this section, we describe the outline of our construction of Nash equilibria, which are all public equilibria. The public strategy of player i is based on a burning-money process $D_i = (D_i^n)_{n \in \mathbf{N}}$, and for every public strategy of length $n-1$, it assigns two parameters: (a) the one-shot mixed action α_i^n to play at stage n and (b) the probability p_i^n to monitor player j at that stage. The monitoring probability p_i^n takes one of three possible values:

- The value $p_i^n = 1$. Here player i is required to burn money, which takes place when his debt exceeds c_i.

When player i's debt is below c_i,

- The value $p_i^n = 0$ when player j plays a best response and need not be monitored.
- The value $p_i^n = p_i$, where p_i is some fixed positive but low constant that satisfies (3). This happens when player j is not playing a best response, hence has to be deterred from deviating.

In principle, the decision whether to monitor may be correlated with the player's action. In our construction, however, the random variables α_i^n and p_i^n are independent, conditional on the current history of length $n-1$.

To facilitate the description of the strategy, we introduce a real-valued process $x_i = (x_i^n)_{n \in \mathbf{N}}$. The quantity x_i^n is the discounted value of the future stream of payoffs starting at stage n, including monitoring fees at stages $m \geq n$, where $p_i^m < 1$. Player i's debt at stage n, D_i^n, indicates the debt of player i at stage n, to be paid by monitoring fees in stages $m \geq n$ in which $p_i^m = 1$.

The actual continuation payoff in the repeated game following the public history h^{n-1} of length $n-1$ under the strategy pair $\sigma = (\sigma_1, \sigma_2)$ is therefore

$$U(\sigma | h^{n-1}) = x^n - D^n. \tag{4}$$

The process $D_i = (D_i^n)_{n \in \mathbf{N}}$ indicates the amount of money player i should burn. We thus require the following condition.

Condition C1. We have $p_i^n = 1$ whenever $D_i^n \geq c_i$, for each player i and every stage n.

Condition C1 means that whenever D_i^n exceeds c_i, player i should burn money by monitoring the other player.

The process $(\alpha_i^n, p_i^n)_{1=1,2;n=1,2,\ldots}$ induces a public equilibrium if for every stage n and player i, the following Conditions (C2)–(C6) are satisfied:

Condition C2. We have $x_i^n - D_i^n \geq v_i$.

Condition C2 ensures that the payoff of each player along the play is individually rational: it is at least his minmax value.

Condition C3. When player j does not play a best response at α^n, for every action $a_j \notin \text{supp}(\alpha_j^n)$,

$$p_i^n > \frac{2(1-r^\Delta)}{r^\Delta(\mathbb{E}_{\alpha_i^n, a_j}[x_i^{n+1} - D_i^{n+1}] - v_i)},$$

where $\mathbb{E}_{\alpha_i^n, a_j}[\cdot]$ denotes the expected value when, at stage n, player i plays the mixed action α_i^n and player j plays the pure action a_j.

As discussed in Section 5.2 (see (3)), Condition C3 ensures that, provided an observed deviation triggers an indefinite punishment at the minmax level, player j cannot profit by deviating to an action that is not in the support of α_j^n.

Because $G(r, c, \Delta)$ is a discounted game, a pair of public strategies is a Nash equilibrium if the behavior of the players following every public history that occurs with positive probability is an equilibrium in the static game in which the payoffs consist of the actual stage payoff plus the continuation payoff (the one induced by σ). The next conditions take care of the incentive compatible constraints associated with this static game.

Denote by I_i^n the event in which player i monitors player j at stage n. Denote by $\mathbb{E}_{a_i, NO_i, \alpha_j^n, p_j^n}[\cdot]$ the expectation operator when at stage n player i plays a_i, he does not monitor (NO_i), while player j plays α_j^n and monitors with probability p_j^n. The notation $\mathbb{E}_{a_i, O_i, \alpha_j^n, p_j^n}[\cdot]$ receives an analog interpretation with the difference that here player i does monitor at stage n.

Condition C4. If $0 < p_i^n < 1$, then for every action $a_i \in \text{supp}(\alpha_i^n)$,

$$\begin{aligned} x_i^n - D_i^n &= (1-r^\Delta)u_i(a_i, \alpha_j^n) + \mathbb{E}_{a_i, NO_i, \alpha_j^n, p_j^n}\big[r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big] \\ &= (1-r^\Delta)u_i(a_i, \alpha_j^n) + \mathbb{E}_{a_i, O_i, \alpha_j^n, p_j^n}\big[r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big]. \end{aligned}$$

Condition C5. If $p_i^n = 0$, then for every action $a_i \in \text{supp}(\alpha_i^n)$,

$$\begin{aligned} x_i^n - D_i^n &= (1-r^\Delta)u_i(a_i, \alpha_j^n) + \mathbb{E}_{a_i, NO_i, \alpha_j^n, p_j^n}\big[r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big] \\ &\geq (1-r^\Delta)u_i(a_i, \alpha_j^n) + \mathbb{E}_{a_i, O_i, \alpha_j^n, p_j^n}\big[r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big]. \end{aligned}$$

Condition C6. If $p_i^n = 1$, then for every action $a_i \in \text{supp}(\alpha_i^n)$,

$$\begin{aligned} x_i^n - D_i^n &= (1-r^\Delta)u_i(a_i, \alpha_j^n) + \mathbb{E}_{a_i, O_i, \alpha_j^n, p_j^n}\big[r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big] \\ &\geq (1-r^\Delta)u_i(a_i, \alpha_j^n) + \mathbb{E}_{a_i, NO_i, \alpha_j^n, p_j^n}\big[r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big]. \end{aligned}$$

Conditions C4–C6 guarantee that no player can profit by an undetectable deviation. Condition C4 states that in case an inspection has a nontrivial probability, all actions in the support of α_i^n guarantee the same payoff, both when an inspection takes place and when it does not. Condition C5 states that in case an inspection occurs with probability 0, all actions in the support of α_i^n guarantee the same payoff if an inspection does not take place and guarantee a lower payoff if an inspection takes place. Thus, there is no incentive to monitor when the probability of monitoring is 0. Similarly, Condition C6 states that in case an inspection occurs with probability 1, all actions in the support of α_i^n guarantee the same payoffs if an inspection takes place, and a lower one if inspection does not take place.[4]

Conditions C4–C6 imply that

$$x_i^n - D_i^n = \mathbb{E}_{p^n,\alpha^n}\big[(1-r^\Delta)u_i(a^n) + r^\Delta(x_i^{n+1} - D_i^{n+1}) - c_i \cdot \mathbf{1}_{I_i^n}\big], \tag{5}$$

which guarantees that (4) holds. Indeed, using (5) recursively one obtains (compare with (1))

$$x_i^N - D_i^N = \mathbb{E}_\sigma\left[\sum_{n=N}^{\infty}(1-r^\Delta)r^{(n-1)\Delta}u_i(a^n) - c\sum_{n=N}^{\infty} r^{(n-1)\Delta}\cdot \mathbf{1}_{I_i^n}\,\middle|\,\mathcal{F}^{n-1}\right], \tag{6}$$

where $\mathbf{1}_{I_i^n}$ is the indicator of the event I_i^n that player i monitors at stage n. The right-hand side of (6) is player i's payoff in the repeated game, starting at stage N.

5.4 *Monitoring to deter deviations*

To better explain the idea of our construction, we start with a simple case in which monitoring is performed for one purpose: to deter deviations. In particular, a burning-money process is unnecessary in this case. In Section 5.5, we handle the general case, which requires the use of burning-money processes.

Suppose that there are two mixed-action pairs $\beta, \gamma \in \Delta(A_1) \times \Delta(A_2)$ that satisfy (see Figure 4) (a) $u_1(\beta) < u_1(\gamma)$, (b) $u_2(\beta) > u_2(\gamma)$, (c) at β, player 1 plays a best response while player 2 is indifferent, and (d) at γ, player 2 plays a best response while player 1 is indifferent. Roughly speaking, we show that any point in the line segment between $u(\beta)$ and $u(\gamma)$ is an equilibrium point.

For every $\eta > 0$, let J_η be the line segment that connects the points $u(\beta) + (\eta, -2\eta)$ and $u(\gamma) + (-2\eta, \eta)$ (see Figure 4).

Assume that the parameters η, r, Δ, c_1, and c_2 satisfy the following smallness conditions: *Condition A1.* $\eta > \frac{1-r^\Delta}{r^\Delta}$; *Condition A2.* $c_1 \le \frac{(u_1(\gamma)-u_1(\beta)-\eta)(r^\Delta-\frac{1}{2})}{1+r^\Delta}$; *Condition A3.* $c_2 \le \frac{(u_2(\beta)-u_2(\gamma)-\eta)(r^\Delta-\frac{1}{2})}{1+r^\Delta}$; *Condition A4.* $r^\Delta\eta^2 > 2c_i > 1 - r^\Delta$ for $i = 1, 2$. Conditions A1–A4 hold whenever $\Delta \ll c_1, c_2 \ll u_1(\gamma) - u_1(\beta) - \eta, u_2(\beta) - u_2(\gamma) - \eta$.

[4]We could state Conditions C4–C6 more concisely. Condition C5 could be required to hold whenever $p_i^n < 1$ (instead of whenever $p_i^n = 0$) and Condition C6 whenever $p_i^n > 0$ (instead of whenever $p_i^n = 1$). In this case, Condition C4 would be redundant. We prefer to keep the three conditions as above for expositional purposes.

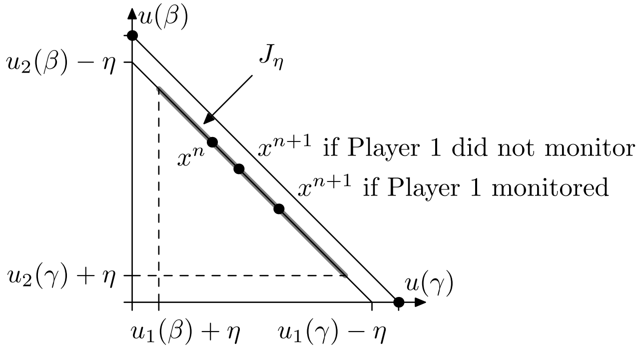

FIGURE 4. The construction in the proof of Lemma 4.

LEMMA 4. *Let $\beta, \gamma \in \Delta(A_1) \times \Delta(A_2)$ be two mixed-action profiles that satisfy the following conditions:*

(i) We have $u_1(\beta) < u_1(\gamma)$ and $u_2(\beta) > u_2(\gamma)$.

(ii) At β, player 1 plays a best response and player 2 is indifferent.

(iii) At γ, player 2 plays a best response and player 1 is indifferent.

Then the set $\mathrm{NE}(r, c_1, c_2, \Delta)$ *contains the line segment J_η, provided that the parameters η, r, c_1, c_2, and Δ satisfy Conditions* A1–A4.

Since player 1 plays a best response at β, we have $v_1 \leq u_1(\beta)$. Similarly, $v_2 \leq u_2(\gamma)$. In particular, all the points on the line segment J_η are feasible and strictly individually rational.

The proof of Lemma 4 is relegated to Section A1 in the Appendix. The key idea of the construction is the following. Any point between $u(\beta)$ and $u(\gamma)$ can be obtained by playing only the mixed actions β and γ throughout the game. When playing β, player 2 may have a profitable deviation, and when playing γ, player 1 may have a profitable deviation. To deter deviations, when playing β, player 1 monitors player 2 with small probability, and when playing γ, player 2 monitors player 1 with small probability. We set the probability of monitoring in such a way that the expected cost of monitoring per stage is η, the distance between the line segment $[u(\beta), u(\gamma)]$ and J_η.

The gist of the construction appears in Figure 4. Let $x^n \in J_\eta$ and suppose furthermore that x^n lies on the upper half of the line segment J_η. To support x^n as the expected payoff from stage n onward, at stage n the players play the mixed action β, and player 1 monitors player 2 with probability p_1. The continuation payoff x^{n+1} depends on whether player 1 monitored player 2: if player 1 monitored player 2, he paid the monitoring cost, and his continuation payoff should be higher than the case when he did not monitor player 2. We thus have to choose two continuation payoffs on J_η that leave the players indifferent. The way this is done and the precise calculations appear in Section A1 in the Appendix. In case x^n lies on the lower half of the line segment J_η, the players play in an analogous fashion: they play the mixed action γ and player 2 monitors player 1 with probability p_2.

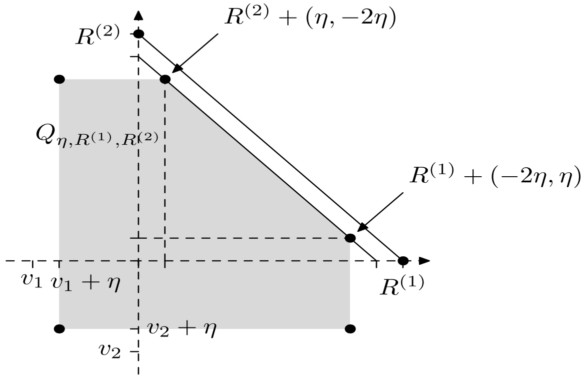

FIGURE 5. The set Q_{η,R_1,R_2}.

Lemma 4 implies in particular that J_η is a subset of $\mathrm{NE}^*(r)$, provided $\eta < \min\{u_1(\gamma) - u_1(\beta), u_2(\beta) - u_2(\gamma)\}$.

5.5 *Monitoring to burn money*

In the next lemma, we relax the conditions of Lemma 4 and extend it in the following ways.

- Condition (i) is violated, that is, $u_1(\beta) \geq u_1(\gamma)$ or $u_2(\beta) \leq u_2(\gamma)$.
- Condition (2) or (3) is violated: player 1 need not be indifferent at γ and player 2 need not be indifferent at β.
- Not only the points on the line segment J_η are equilibrium payoffs, but also all individually rational points dominated by points on J_η.

Note that we keep the conditions that player 1 plays a best response at β and that player 2 plays a best response at γ, which are the only crucial requirements for our construction.

For every two points $R^{(1)}, R^{(2)} \in \mathbb{R}^2$ that satisfy $R_1^{(1)} > R_1^{(2)} \geq v_1$ and $R_2^{(2)} > R_2^{(1)} \geq v_2$, and for every $\eta > 0$ that satisfies $\eta < \frac{R_i^{(3-i)} - v_i}{2}$ and $\eta < \frac{R_i^{(i)} - R_i^{(3-i)}}{3}$ for $i = 1, 2$, let $Q_{\eta,R^{(1)},R^{(2)}}$ be the pentagon whose extreme points are $(v_1 + \eta, v_2 + \eta)$, $(v_1 + \eta, R_2^{(2)} - 2\eta)$, $(R_1^{(1)} - 2\eta, v_2 + \eta)$, $(R_1^{(2)} + \eta, R_2^{(2)} - 2\eta)$, and $(R_1^{(1)} - 2\eta, R_2^{(1)} + \eta)$ (see Figure 5). It is convenient to phrase the result in terms of asymptotic sets of Nash equilibrium payoffs.

LEMMA 5. *Let $\beta = (\beta_1, \beta_2)$ and $\gamma = (\gamma_1, \gamma_2)$ be two mixed-action pairs and let $R^{(1)}, R^{(2)} \in \mathbb{R}^2$ be such that the following conditions hold:*

(i) We have $u_1(\beta) > R_1^{(2)} > v_1$ and $u_2(\gamma) > R_2^{(1)} > v_2$.

(ii) Player 1 plays a best response at β and $u_2(\beta_1, a_2) \geq R_2^{(2)}$ for every action $a_2 \in$ $\mathrm{supp}(\beta_2)$.

(iii) *Player 2 plays a best response at* γ *and* $u_1(a_1, \gamma_2) \geq R_1^{(1)}$ *for every action* $a_1 \in \operatorname{supp}(\gamma_1)$.

Then the set $Q_{0,R^{(1)},R^{(2)}}$ *is an asymptotic set of Nash equilibrium payoffs.*

It is sufficient to prove that for every $\eta > 0$, the set $Q_{\eta,R^{(1)},R^{(2)}}$ is a subset of $\text{NE}(r, c_1, c_2, \Delta)$, provided that c_1, c_2, and Δ are sufficiently small. The proof of Lemma 5 is based on the construction employed in the proof of Lemma 4: the players play either β or γ, and whenever a player is supposed to play a mixed action that is not a best response, he is monitored with a small probability. To prove the lemma, we use burning-money processes. These processes have several roles:

- To implement an equilibrium payoff ξ in the interior of $Q_{\eta,R^{(1)},R^{(2)}}$, we choose a point x on the line segment $[R^{(2)} + (\eta, -2\eta), R^{(1)} + (-2\eta, \eta)]$ that dominates ξ. We implement x and use burning-money processes with initial values $(x_i - \xi_i)_{i=1,2}$. This way each player i obtains x_i, burns $x_i - \xi_i$, and is left with ξ_i.
- When a player is not indifferent between the pure actions in the mixed action he is supposed to play, we use a burning-money process to make him indifferent between the pure actions involved. Suppose for example that player 1 is not indifferent at γ. Whenever player 1 plays γ and is being monitored, his debt increases by $\frac{u_1(a_1,\gamma_2)-R_1^{(1)}}{p_2} \geq 0$, where a_1 is the actual action that he played. This quantity is player 1's observed excess over $R_1^{(1)}$ divided by the per-stage probability that he is monitored. This way, player 1's debt increases by a quantity that makes him indifferent between the pure actions in $\operatorname{supp}(\gamma_1)$.
- To accommodate the case that $u_1(\beta) \geq u_1(\gamma)$ or $u_2(\beta) \leq u_2(\gamma)$, we apply the construction of Lemma 4, *assuming* that the players' payoff when they play β (resp. γ) is $R^{(2)}$ (resp. $R^{(1)}$) rather than $u(\beta)$ (resp. $u(\gamma)$). Since the payoff is in fact $u(\beta)$ (resp. $u(\gamma)$), at every stage in which the players play β (resp. γ), we add $(1 - r^\Delta)(u(\beta) - R^{(2)})$ (resp. $(1 - r^\Delta)(u(\gamma) - R^{(1)})$) to the players' debts.

The details of the proof appear in Section A2.

By choosing β to be the mixed action in which the maximum in the definition of M_2 is attained and choosing γ to be the mixed action in which the maximum in the definition of M_1 is attained, we obtain that $R_1^{(1)}$ (resp. $R_2^{(2)}$) can be assumed to equal to M_1 (resp. M_2). From Lemma 5 we therefore obtain the following result.

Corollary 1. *Assume that* $M_1 > v_1$ *and* $M_2 > v_2$. *Let* β *be a mixed-action profile in which* β_2 *is a best response and* $\min_{a_1 \in \operatorname{supp}(\beta_1)} u_1(a_1, \beta_2) = M_1$, *and let* γ *be a mixed-action profile in which* γ_1 *is a best response and* $\min_{a_2 \in \operatorname{supp}(\gamma_2)} u_2(\gamma_1, a_2) = M_2$. *Then the pentagon* Q^* *whose extreme points are* (v_1, v_2), (v_1, M_2), $(u_1(\beta), M_2)$, $(M_1, u_2(\gamma))$, *and* (M_1, v_2) *is an asymptotic set of Nash equilibrium payoffs.*

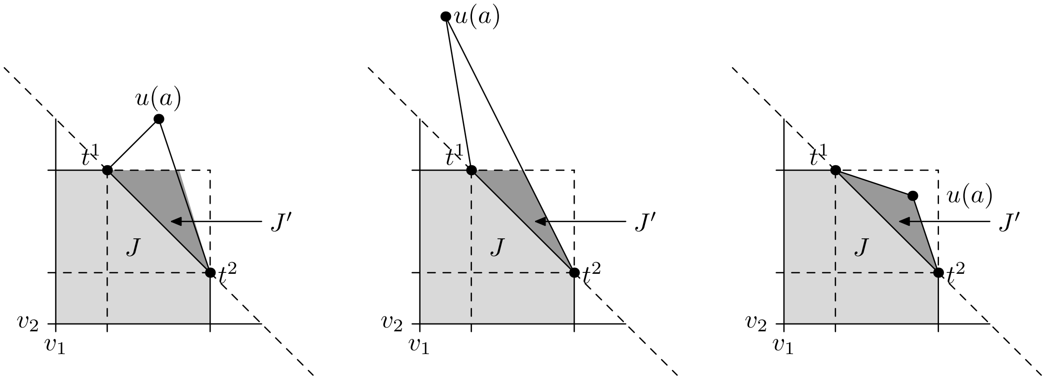

FIGURE 6. Three cases of $u(a)$ in Lemma 6 with $t^1_2 = M_2$ and $t^2_1 = M_1$.

5.6 *Equilibria in which both players monitor each other*

The construction in the previous section uses action pairs in which one player plays a best response and therefore need not be monitored. In the present section, we use action pairs in which neither player plays a best response. Consequently, there is a positive probability that both players monitor each other.

For every two points $x, y \in \mathbb{R}^2$, denote by $\langle\langle x; y\rangle\rangle$ the rectangle with extreme points x, (x_1, y_2), (y_1, x_2), and y.

LEMMA 6. *Let J be an asymptotic set of Nash equilibrium payoffs, let t^1 and t^2 be two points in J such that the slope of the line segment $[t^1, t^2]$ is negative and such that the rectangles $\langle\langle v; t^1\rangle\rangle$ and $\langle\langle v; t^2\rangle\rangle$ are subsets of J, and let a be an action profile such that $u(a)$ is not dominated by any point on the line that passes through t^1 and t^2 (see Figure 6). The intersection of F^M and the convex hull of J and $u(a)$ is an asymptotic set of Nash equilibrium payoffs.*

Denote by J' the intersection of F^M and the triangle whose extreme points are t^1, t^2, and $u(a)$. The idea is to define for every payoff vector x in the triangle J', a one-shot game in which (a) each player has two actions, to monitor or not to monitor, and (b) the payoff is $(1 - r^\Delta)u(a) + r^\Delta y$, where y is a continuation payoff that depends on the actions of the players and is in $J \cup J'$. We choose the continuation payoffs in such a way that there is an equilibrium in the one-shot game in which (a) the expected payoff is x and (b) both players monitor each other with a probability that satisfies (3). Since the game is discounted, an iterative use of this construction yields for any point $y \in J'$ an equilibrium of the repeated game with payoff y. The details are elaborated in Section A3 in the Appendix.

We argue that an iterative application of Lemma 6 implies that $F^M \subseteq \mathrm{NE}^*(r)$. Let γ^* be a mixed-strategy pair in which M_1 is attained; that is, player 2 plays a best response at γ^* and $\min_{a_1 \in \mathrm{supp}(\gamma^*_1)} u_1(a_1, \gamma^*_2) = M_1$. Similarly, let β^* be the analogous mixed-action pair with respect to M_2. Denote $t_1 := u_1(\beta^*)$ and $t_2 := u_2(\gamma^*)$, and let J be the pentagon whose extreme points are (v_1, v_2), (v_1, M_2), (M_1, v_2), (t_1, M_2), and (M_1, t_2)

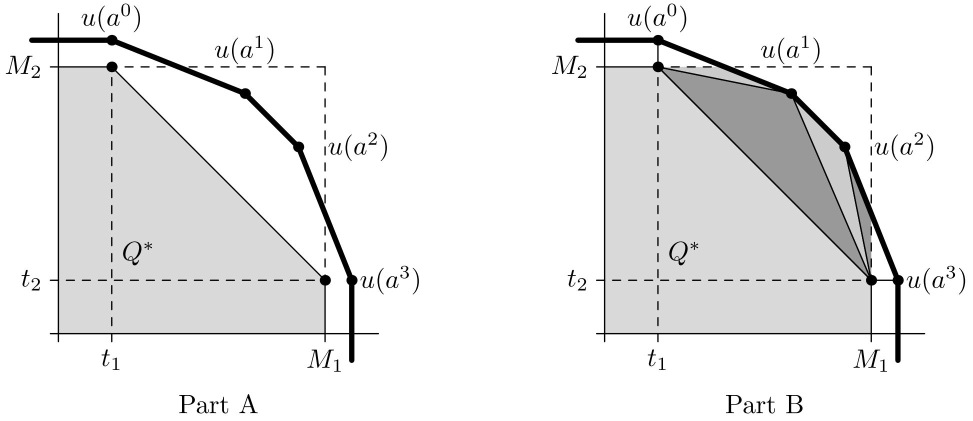

Figure 7. There are payoff vectors $u(a)$ in the Pareto frontier of F^M. The Pareto frontier of F^M is denoted by the dark line.

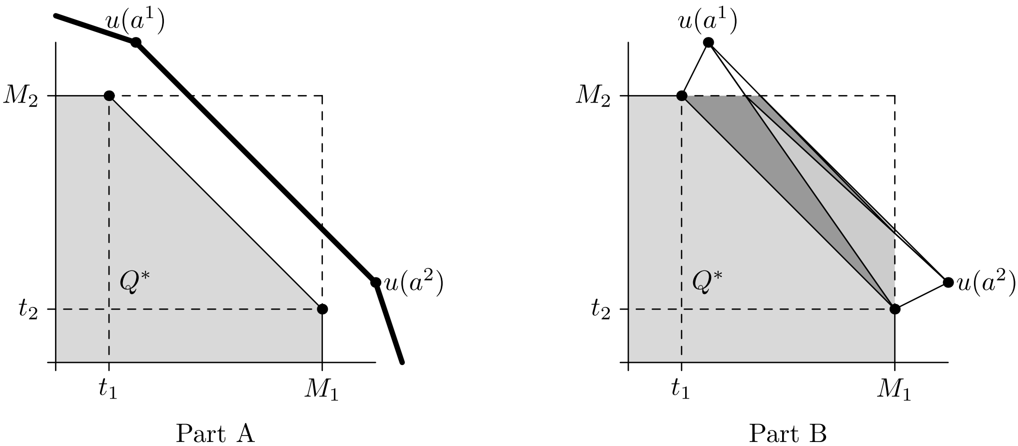

Figure 8. No payoff vectors $u(a)$ in the Pareto frontier of F^M.

(see Figure 7(A)). By Lemma 5 with $t^1 = (t_1, M_2)$ and $t^2 = (m_1, t^2)$, the pentagon J is an asymptotic set of Nash equilibrium payoffs.

We distinguish between two cases, according to whether there are payoff vectors $u(a)$ in the Pareto frontier of F^M. Consider first the case that appears in Figure 7(A), in which the Pareto frontier of F^M contains the points $u(a^1)$ and $u(a^2)$. Apply Lemma 6 to J and $u(a^1)$ to prove that the convex hull of J and $u(a^1)$ is an asymptotic set of Nash equilibrium payoffs (see Figure 7(B)). Apply next Lemma 6 to $u(a^2)$ and to the Pareto frontier of the convex hull of J and $u(a^1)$ to prove that the convex hull of J, $u(a^1)$, and $u(a^2)$ is an asymptotic set of Nash equilibrium payoffs (see Figure 7(B)). To conclude that F^M is an asymptotic set of Nash equilibrium payoffs, apply Lemma 6 to the Pareto frontier of the resulting set and to $u(a^0)$ and $u(a^3)$.

Consider now the case illustrated in Figure 8(A), in which there are no payoff vectors $u(a)$ in the Pareto frontier of F^M. Note that each application of Lemma 6 generates a larger pentagon. We apply Lemma 6 to $u(a^1)$ and to $u(a^2)$ alternately, each time with the previously generated pentagon (see Figure 8(B)). The sequence of pentagons thus generated converges to the intersection of F^M and the convex hull of J, $u(a^1)$, and $u(a^2)$. This way we show that the intersection of F^M and the convex hull of J, $u(a^1)$, and $u(a^2)$ is an asymptotic set of Nash equilibrium payoffs.

The general observation, which allows us to conclude that an iterative application of Lemma 6 completes the construction, is the next lemma. Recall that J is the pentagon whose extreme points are (v_1, v_2), (v_1, M_2), (M_1, v_2), (t_1, M_2), and (M_1, t_2).

Lemma 7. *Let X be the smallest subset of F^M that satisfies the following properties: (a) X contains Q^* (see Corollary 1) and (b) for every two points t^1 and t^2 in X such that the slope of the line segment $[t^1, t^2]$ is negative and such that the rectangles $\langle\langle v; t^1\rangle\rangle$ and $\langle\langle v; t^2\rangle\rangle$ are subsets of X, and for every action profile a such that $u(a)$ is not dominated by any point on the line connecting t^1 and t^2, the set X contains the intersection of F^M and the triangle $\langle t^1, t^2, u(a)\rangle$. Then X is the intersection of F^M and the convex hull of Q^* and the set $\{u(a) \colon a \in A, a \geq v\}$.*

5.7 *Public perfect equilibria: Showing that* $\mathrm{NE}^*(r) \subseteq \mathrm{PPE}^*(r)$

In the construction of Nash equilibria, we use threats of punishment. To show that $\mathrm{NE}^*(r) \subseteq \mathrm{PPE}^*(r)$, we observe that in our construction the continuation payoffs of each player i are at least $v_i + \eta$, and therefore it suffices to show that the vector $\xi^* := (v_1 + \eta, v_2 + \eta)$ is a public perfect equilibrium payoff, provided η is sufficiently close to 0. This way, every Nash equilibrium can be transformed into a public perfect equilibrium, by having the players switch to a PPE that implements ξ^* instead of punishing a deviator at his minmax value.

Lemma 8. *Suppose that $\eta > 0$ satisfies $\eta < \min\{M_1 - v_1, M_2 - v_2\}$. Then $(v_1 + \eta, v_2 + \eta)$ is a PPE payoff.*

As is common in the literature, the implementation of a credible punishment is accomplished by having the players lower their payoffs for a fixed number of stages before returning to the equilibrium path. This is attained by having each player play a minmax strategy against the other player in the base game while monitoring the other player for a fixed number of stages. Monitoring the other player lowers one's own payof and is observable. If at some stage a player fails to monitor the other, the counter that counts the number of stages in which players monitored each other is reset. During these stages each player plays a minmax strategy against the other player, hence the expected payoff of each player i during this phase is below v_i. Consequently, lengthening this phase would not be profitable. To ensure that players do not have incentives to deviate during these stages by playing a mixed action that is not their minmax strategy, we note that once the phase of lowering payoffs ends, each player knows the payoffs that the other player has accumulated since the counter was last reset, and thus, by using burning-money processes in future stages, we can render the players indifferent between future plays that are consistent with the equilibrium play. For more details, see Section A7 in the Appendix.

6. Comments and open problems

6.1 *Repeated games in which the one-shot game has no equilibrium*

Similarly to the folk theorem, our construction does not require the existence of an equilibrium in the base game. Indeed, the proof is valid as soon as M_1 and M_2 are well defined and $M_i > v_i$ for each $i \in \{1, 2\}$. Such a case occurs, for example, in the following game with infinitely many actions. This game is the well known game of picking the higher natural number, supplemented with a punishment action (the action P), which gives a bad payoff to both players.

- There are two players with identical action sets: $A_1 = A_2 = \{P, 1, 2, \ldots\}$.
- The payoff function is

$$u(a_1, a_2) = \begin{cases} (1, 0) & \text{if } 1 \le a_2 < a_1, \\ (0, 1) & \text{if } 1 \le a_1 < a_2, \\ \left(\frac{1}{2}, \frac{1}{2}\right) & \text{if } 1 \le a_1 = a_2, \\ (-1, -1) & \text{if } a_1 = P, a_2 \ge 1 \text{ or } a_1 \ge 1, a_2 = P, \\ (-2, -2) & \text{if } a_1 = a_2 = P. \end{cases}$$

In this case, $M_1 = M_2 = 0$, $v_1 = v_2 = -1$, and $\mathrm{PPE}^*(r) = \mathrm{NE}^*(r) = [-1, 0]^2$.

6.2 *More than two players*

Two difficulties arise when more than two players are present. First, the players face a coordination problem: if player i monitors player j and finds out that player j deviated, to effectively punish the deviator, player i has to make the fact that the deviation occurred common knowledge, and it is not clear how this task can be achieved. This difficulty does not arise in cases in which the action played by a monitored player is observed by *all* other players and not just by the player who paid the monitoring fee. Indeed, in this case all players share the same information on each other, and they all know when a deviator should be punished.

Second, one has to adapt the equilibrium construction to more than two players. While the construction in the proof of Lemma 5 can be adapted to games with any number of players (assuming monitored actions become common knowledge), we do not know whether the construction in the proof of Lemma 6 generalizes to more than two players.

6.3 *Cheap talk*

Another interesting issue that arises when more than two players are involved concerns cheap talk. Suppose that the players can costlessly send messages at the beginning of every stage. When only two players are present, the characterization of PPE payoffs and NE payoffs does not change. However, when three players or more are present, new coordination schemes become possible. For example, player 1 may ask player 2 to play a specific action and ask player 3 to monitor it and to report the action he observed. This kind of communication might expand the set of equilibrium payoffs. We do not know whether this is the case and, if so, what the new set of equilibrium payoffs is.

6.4 *Monitoring is imperfect and public*

We assumed that the act of monitoring is common knowledge. This assumption is crucial for our results: if monitoring goes unnoticed by the other party, there is no way to compensate a player for monitoring the other player.

Consider now a model in which monitoring is imperfect yet public: at the end of every period, the players observe two public signals, which indicate whether each of them monitored the other at that stage. If player i did not monitor player j, then the public signal indicates that no monitoring was done by player i; if player i monitored player j, then with some known probability, the signal indicates that monitoring was done by player i, and with the remaining probability, it indicates that monitoring was not done by player i. Our results carry over to this model: the only differences in the construction of the equilibrium strategies are that (a) any unobserved monitoring is ignored and (b) the effective monitoring cost is updated to reflect the fact that sometimes monitoring is ignored.

6.5 *Different discount factors*

In this paper, we assume that the discount factors of the two players coincide. We leave open the question of characterizing the equilibrium sets in case the players differ in their time preferences (see Lehrer and Pauzner (1999) for a characterization of the set of subgame-perfect equilibrium payoffs in repeated games without monitoring cost and with different discount factors).

6.6 *The method of fudenberg and levine*

A natural question arises regarding whether the general method of Fudenberg and Levine (1994) applies to our model. Since the monitoring cost here is "infinitely" higher than the stage payoff, we do not know whether their method could be applied.

References

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [91]

Ben-Porath, Elchanan and Michael Kahneman (2003), "Communication in repeated games with costly monitoring." *Games and Economic Behavior*, 44, 227–250. [91]

Compte, Olivier (1998), "Communication in repeated games with imperfect private monitoring." *Econometrica*, 66, 597–626. [91]

Flesch, János and Andrés Perea (2009), "Repeated games with voluntary information purchase." *Games and Economic Behavior*, 66, 126–145. [91]

Fudenberg, Drew, David M. Kreps, and Eric S. Maskin (1990), "Repeated games with long-run and short-run players." *The Review of Economic Studies*, 57, 555–573. [87, 89, 90]

Fudenberg, Drew, David Levine, and Eric Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62, 997–1039. [91]

Fudenberg, Drew and David K. Levine (1994), "Efficiency and observability in games with long-run and short-run players." *Journal of Economic Theory*, 62, 103–135. [91, 112]

Hörner, Johannes, Takuo Sugaya, Satoru Takahashi, and Nicolas Vieille (2011), "Recursive methods in discounted stochastic games: An algorithm for $\delta \to 1$ and a folk theorem." *Econometrica*, 79, 1277–1318. [91]

Kandori, Michihiro and Hitoshi Matsushima (1998), "Private observation, communication and collusion." *Econometrica*, 66, 627–652. [91]

Lehrer, Ehud (1989), "Lower equilibrium payoffs in repeated games with non-observable actions." *International Journal of Game Theory*, 18, 57–89. [91]

Lehrer, Ehud (1990), "Nash equilibria of n-player repeated games with semi-standard information." *International Journal of Game Theory*, 19, 191–217. [91]

Lehrer, Ehud (1991), "Internal correlation in repeated games." *International Journal of Game Theory*, 19, 431–456. [91]

Lehrer, Ehud (1992), "Two player repeated games with non-observable actions and observable payoffs." *Mathematics of Operations Research*, 17, 200–224. [91]

Lehrer, Ehud and Ady Pauzner (1999), "Repeated games with differential time preferences." *Econometrica*, 67, 393–412. [112]

Lipman, Barton L. and Ruqu Wang (2009), "Switching costs in infinitely repeated games." *Games and Economics Behavior*, 66, 292–314. [91]

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York. [90]

Mertens, Jean-François, Sylvain Sorin, and Shmuel Zamir (2016), *Repeated Games*. Cambridge University Press, Cambridge. [90]

Miyagawa, Eiichi, Yasuyuki Miyahara, and Tadashi Sekiguchi (2008), "The folk theorem for repeated games with observation costs." *Journal of Economic Theory*, 139, 192–221. [91]

Pearce, David G. (1992), "Repeated games: Cooperation and rationality." In *Advances in Economic Theory: Volume 1: Sixth World Congress* (Jean-Jacques Laffont, ed.), 132–174, Cambridge University Press, New York. [90]

Co-editor George J. Mailath handled this manuscript.

Manuscript received 6 September, 2016; final version accepted 5 May, 2017; available online 14 June, 2017.