The Korean Economic Review
Volume 35, Number 1, Winter 2019, 109-123.

Repeated Games with Asymptotically Finite Horizon and Imperfect Public Monitoring

Yves Guéron*

We consider a two-player infinitely repeated game with asymptotically finite horizons: discount factors converge to zero over time. The stage-game has a continuum of actions and a unique and interior Nash equilibrium. It is known that when players perfectly observe each other's actions, cooperation can be achieved and equilibrium payoffs can be strictly higher than the stage-game equilibrium payoff. We show that introducing an arbitrarily small amount of smooth noise in the monitoring makes cooperation impossible and players play the static Nash equilibrium of the stage-game forever.

JEL Classification: C72, C73
Keywords: Dynamic Games, Asymptotically Finite Horizon, Folk-theorem, Imperfect Monitoring

I. Introduction

Games that have a unique stage-game Nash equilibrium have the undesirable property that when repeated a finite number of times, the only subgame-perfect equilibrium is the repetition of the stage-game Nash equilibrium. That is, there are no intertemporal incentives at play and players often end up along an inefficient path, such as in the finitely repeated prisoner's dilemma.

Infinitely repeated games provide a satisfying answer to this problem. An infinite number of repetitions gives rise to new intertemporal incentives as there is no longer a last period from which incentives unravel. The folk theorem then tells us that every feasible and strictly individually rational payoff can be supported as the outcome of a subgame-perfect equilibrium.

Received: Nov. 22, 2017. Revised: Jan. 26, 2018. Accepted: July 6, 2018.

* Assistant Professor, Department of Economics, Seoul National University, Seoul 08826, Korea. ygueron@snu.ac.kr. This work was supported by Research Resettlement Fund for the new faculty of SNU and by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2014S1A5A8017310). I thank George Mailath for mentioning the Bernheim and Dasgupta paper.

With constant discounting, infinitely repeated games are stationary. In particular, the expected duration of the continuation game is the same at any time.[1] One might however argue that even though a game may not have deterministic ending, its expected length should decrease as time passes. For example, two competitors facing a declining demand might expect the proba-bility of interacting to decline over time due to increasing risks of shutdowns. Accordingly, Bernheim and Dasgupta (1995) (hereafter BD) introduce repeated games with asymptotically finite horizons. In each period, there is a strictly positive probability that the game continues to the next period. However this probability converges to zero.

BD study a stage game in which action sets are compact and continuous and which has a unique interior and locally inefficient Nash equilibrium. The game is infinitely repeated and δ_t, the discount factor applied from period t to period $t+1$, is such that $\delta_t > 0$ for all t and $\lim_{t\to\infty} \delta_t = 0$. They show that, provided discount factors do not converge to zero too fast, it is possible to have subgame-perfect equilibria in which in each period players receive a payoff strictly higher than the stage-game Nash equilibrium payoff. However, they also show that in any such equilibria, action profiles will converge to the stage-game Nash equilibrium. Nonetheless, this can occur quite late in the game, and thus have little impact on period-zero payoffs, such that full efficiency may be approximatively achieved.

In this paper we show that when introducing an arbitrarily small amount of smooth public noise, all the non-degenerate equilibria of BD break down, and the only equilibrium of the dynamic game is the infinite repetition of inefficient stage-game Nash equilibrium.

The result is similar to Guéron (2015), who studies dynamic contribution games with imperfect monitoring and shows that the introduction of a small amount of noise causes a complete breakdown in cooperation. However the stage-game considered is quite different: in this paper, the stage-game has a unique interior Nash equilibrium, whereas in Guéron (2015) the stage-game is dominant-solvable.

II. Perfect Monitoring

In this section we briefly present the model of Bernheim and Dasgupta (1995) and discuss their main results.

2.1. The Stage Game

There are two players,[2] $i = 1, 2$. Let A_i denote player i's action set and

[1] If the discount factor is interpreted as the probability that the interaction continues in the next period.

[2] This is only to simplify exposition, but all results extend to the case of n players.

$u_i : A_i \times A_j \to \mathbb{R}$ denote his utility function. We summarize below all assumptions on the stage game:

Assumption 1. *$A_i := [\underline{a}_i, \overline{a}_i]$ is a compact subset of $\mathbb{R}$. There is a unique stage-game Nash equilibrium a^N which is contained in the interior of $A := A_1 \times A_2$. Each payoff function $u_i : A \to \mathbb{R}$ is twice continuously differentiable with respect to a in a neighborhood of a^N and strictly quasi-concave in a_i. Each player's best reply function $\phi_i : A_j \to A_i$ is continuously differentiable in a neighborhood of a^N. The Jacobian matrix of partial derivatives Du has full rank at the stage-game Nash equilibrium.*[3,4]

Assumption 2. *Utility functions u_i are three times continuously differentiable in a neighborhood of a^N, $D^2 u_i(a^N)$ is negative definite, $i = 1, 2$ and $D\phi(a^N) - I$ is non-singular, where $D^2 u_i$ is the Hessian matrix of u_i and $D\phi$ the Jacobian matrix of the best response function.*

Remark 1. While BD assume that utility functions are three times continuously differentiable, we only use the fact that they are twice continuously differentiable in what follows. Moreover we will not require that $D^2 u_i(a^N)$ is negative definite but only that $\partial^2 u_i / \partial a_i^2 < 0$ at the Nash equilibrium.

Example 1. A Cournot duopoly with linear inverse demand and constant marginal cost satisfies Assumptions 1 and 2. Let $A_1 = A_2 = [0, \overline{q}]$, where $\overline{q}$ is sufficiently high, and let $u_i(q_i, q_j) = \max\{0, q_i(1 - q_i - q_j)\}$, $i = 1, 2$, $j \neq i$.

2.2. The Dynamic Structure

Time is discrete and the game is played infinitely many times: $t = 0, 1, \ldots$. Players share a common sequence of discount factors $(\delta_t)_{t \geq 0}$, where $\delta_t \in (0,1)$ is the discount rate from period t to $t+1$. The game has an asymptotically finite horizon in the sense that $\lim_{t\to\infty} \delta_t = 0$.

Payoff streams are evaluated using unnormalized discounting. The payoff to player i from a stream of payoffs $(u_i^0, u_i^1, \ldots)$, evaluated in period k, $k \geq 0$ is:[5]

---

[3] Because u_i is strictly quasi-concave in a_i, each player i has a unique best response to any action a_j.

[4] In the case of two players, the Jacobian matrix evaluated at the Nash equilibrium will be $Du(a^N) = \begin{pmatrix} 0 & \frac{\partial u_1}{\partial a_2}(a^N) \\ \frac{\partial u_2}{\partial a_1}(a^N) & 0 \end{pmatrix}$. The full rank assumption then reduces to the assumption that at the Nash equilibrium $\frac{\partial u_1}{\partial a_2}$ and $\frac{\partial u_2}{\partial a_1}$ are non-zero.

[5] With the convention that $\prod_{\tau=k}^{k-1} \delta_\tau = 1$.

$$\sum_{t=k}^{\infty}\left\{\prod_{\tau=k}^{t-1}\delta_\tau\right\}u_t = u_k + \delta_k u_{k+1} + \delta_k\delta_{k+1}u_{k+2} + \delta_k\delta_{k+1}\delta_{k+2}u_{k+3} + \dots$$

Players discount the future in the sense that $\sum_{t=0}^{\infty}\{\prod_{\tau=0}^{t-1}\delta_\tau\} < \infty$. For ease of notation we define $\beta_k^t := \prod_{\tau=k}^{t-1}\delta_\tau$ to be the rate at which period t payoffs are discounted in period k, $k \leq t$.

2.3. Cooperation in Repeated Games with Asymptotically Finite Horizons and Perfect Monitoring

In this section we briefly present the main results of BD. First, they find a sufficient condition on the rate of convergence of discount factors to zero to guarantee the existence of a non-degenerate subgame-perfect equilibrium, that is an equilibrium in which players obtain a payoff strictly greater than the stage-game Nash equilibrium payoff in each period. More specifically, the log of discount rates must grow, in absolute value, faster than 2^k:

Assumption 3. *There exist* $c > 0$ *and* $\Lambda > 0$ *such that* $\prod_{k=0}^{\tau-1}\delta_k^{2^{\tau-1-k}} \geq c\Lambda^{2\tau}$, $\tau \geq 1$. *This is equivalent to having* $\lim_{\tau\to\infty}\sum_{k=0}^{\tau-1}\frac{1}{2^{k+1}}\ln(\delta_k) > -\infty$.

Theorem 1 (Bernheim and Dasgupta 1995). *Under Assumptions 1 and 3, there exists a subgame-perfect equilibrium in which players receive a payoff strictly higher than the stage-game NE payoff in each period.*

Furthermore, in any equilibrium, the action profile converges to the unique interior Nash equilibrium of the stage game.

Under the additional regularity of Assumption 2 BD show that assumption 3 is not only sufficient but necessary for the existence of non-degenerate subgame-perfect equilibria.

They also establish a folk theorem. More specifically they prove the existence of a time T^* such that for any T, if discount factors are above a certain threshold for $T+T^*$ periods and decline sufficiently slowly, for any feasible, interior and strictly individually rational payoff v, there is a subgame-perfect equilibrium in which players get v for at least T periods.

III. The Model under Public Monitoring

We now introduce imperfect public monitoring to the model. At the end of each period, and conditional on an action profile a, players only observe a public signal

y drawn from a compact set $Y \subset \mathbb{R}^m (m \geq 1)$, according to a probability measure $\pi(\cdot \mid a)$. (Players do not observe each other's actions.) For any measurable $E \subset Y$ we have:

$$\mathbb{P}(y \in E \mid a) = \int_E \pi(dy \mid a).$$

Our main assumption is that the probability measure π is continuous with respect to action profiles. The idea behind this assumption is that if a player changes his action by a small amount only, the change in the distribution of signals will also be small:

Assumption 4 (Feller continuity). *There exists a constant K such that $|\mathbb{P}(E \mid a_1 + \Delta, a_2) - \mathbb{P}(E \mid a_1, a_2)| \leq K\Delta$ and $|\mathbb{P}(E \mid a_1, a_2) - \mathbb{P}(E \mid a_1, a_2 + \Delta)| \leq K\Delta$ for any measurable set $E \subset Y$.*

We denote player i's realized payoff by u_i^*, which is a function of his current action and the public signal. The ex ante utility function u_i is then the expectation of the ex post payoff:[6]

$$u_i(a) = \int_Y u_i^*(a_i, y)\pi(dy \mid a), \quad \forall a \in A. \tag{1}$$

A public history h^t is a sequence of t public signals: $h^t = (y^0, y^1, \ldots, y^{t-1}) \in Y^t$. The set of all public histories is $\mathcal{H} := \cup_{t \geq 0} Y^t$.

A public behavior strategy σ_i for player i is a measurable function that specifies a probability distribution $\sigma_i(h^t) \in \Delta(A_i)$ after any public history $h^t \in \mathcal{H}$:

$$\sigma_i : \begin{cases} \mathcal{H} & \to \Delta(A_i) \\ h^t & \mapsto \sigma_i(h^t) \end{cases}.$$

The monitoring technology, along with a public strategy profile $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)$,

---

[6] When players receive payoffs at the end of every period, this payoff cannot depend on the other player's action other than through the public signal. This is because the public signal is the only information available regarding the other player's action. This is why the realized payoff function u_i^* is introduced, even though as in most of the literature on repeated games with imperfect monitoring, we work with the ex ante utility function.

Here is nevertheless a simple example of realized payoff. Suppose that the public signal $y = (y_1, y_2)$ is two-dimensional and such that $\mathbb{E}(y_1, y_2) = (q_1, q_2)$. Then we could have $u_i^*(a_i, y) := u_i(a_i, y_j)$, in which case (1) would be satisfied − as long as u_i^* is linear in y_j. For example in the case of a linear inverse demand we could have $u_i^*(a_i, y) = a_i(1 - a_i - y_j)$ and $\mathbb{E}(u_i^*(a_i, y)) = \mathbb{E}(a_i(1 - a_i - y_j)) = a_i(1 - a_i - \mathbb{E}(y_j)) = a_i(1 - a_i - a_j)$.

induce a probability distribution on $\mathcal{H}$ that we denote by $\mathbb{P}_{\boldsymbol{\sigma}}$. Expectations with respect to that probability distribution will be denoted by $\mathbb{E}_{\boldsymbol{\sigma}}$.[7]

Let $V_i(\sigma_i, \sigma_j)$ be the expected payoff of player i from the strategy profile $\boldsymbol{\sigma} = (\sigma_1, \sigma_2)$ and $V_i(\sigma_i, \sigma_j \mid h^\tau)$ be the continuation payoff from $\boldsymbol{\sigma}$ after the public history h^τ :

$$V_i(\sigma_i, \sigma_j) := \mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{t=0}^{\infty} \beta_0^t u_i(a^t)\right],$$

$$V_i(\sigma_i, \sigma_j \mid h^\tau) := \mathbb{E}_{\boldsymbol{\sigma}}\left[\sum_{t=0}^{\infty} \beta_\tau^{\tau+t} u_i(a^{\tau+t}) \mid h^\tau\right].$$

A profile of public behavior strategies (σ_1, σ_2) is a Nash equilibrium if for any $i \in \{1,2\}$ and any strategy σ' we have that $V_i(\sigma_i, \sigma_j) \geq V_i(\sigma', \sigma_j)$.[8]

IV. Breakdown of Cooperation with Public Monitoring

We now show that cooperation breaks down under "smooth" imperfect public monitoring, even for an arbitrarily small amount of noise.

First, as in BD, in equilibrium actions will converge to the unique interior stage-game Nash equilibrium. Hence, for t sufficiently large, small deviations will offer a gain of order two, from the envelope theorem.

Consider now the cost from such a deviation. First, as discount factors converge to zero, intertemporal incentives eventually become weak and future costs have little impact on current utility. Second, as action profiles approach the unique stage-game Nash equilibrium, the cost of being punished by a permanent play of the stage-game Nash equilibrium is low.[9] Finally, because of the Feller continuity assumption, a small deviation will have a low impact on the monitoring, implying a low probability of punishment. Overall, the cost of deviating will be lower than the gain, so that cooperation cannot be sustained:

Theorem 2. *Consider an infinitely repeated game with asymptotically finite horizon and imperfect public monitoring that satisfy assumptions 1, 2, 3 and 4. Then with probability one the only Nash equilibrium in public strategies is the infinite repetition of*

---

[7] For any finite t, we can easily define the probability distribution over Y^t by using the finite product measure implied by the distribution of the public signal. This can be extended to the set of infinite public histories $\mathcal{H}$ thanks to the Kolmogorov extension theorem.

[8] Given that u_i is bounded, V_i is well defined.

[9] This is the only feasible punishment, as action profiles must converge to the unique stage-game Nash equilibrium in any equilibrium.

the unique stage-game Nash equilibrium:[10,11]

$$\mathbb{P}_{\sigma}\{a^t = a^N\} = 1, \forall t \geq 0 .$$

Remark 2 (Assumptions). Before proceeding to the analysis, let us briefly summarize the use of each assumption. Assumption 1, is the main assumption about the structure of stage game and is used throughout. Assumption 2 is an additional regularity assumption about the stage game, and Assumption 3 is the requirement that discount factors do not converge too fast to zero. BD show that under Assumptions 1 and 2, Assumption 3 is necessary and sufficient for the existence of a non-degenerate subgame-perfect equilibrium.

In what follows, in order to prove Theorem 2, we will use Assumptions 1, 2 and 4, which is our main assumption about the continuity of the monitoring technology. Although We do not use Assumption 3, we include it nonetheless in the statement of Theorem 2 because, without this assumption, there cannot be a non-degenerate equilibrium in the repeated game with perfect monitoring.

4.1. Preliminary Results

First, we show in Lemma 1 that in any non-degenerate Nash equilibrium of the dynamic game,[12] for any t, there must be a set of histories of length at least t for which stage-game payoffs are strictly higher than the stage-game Nash equilibrium payoff. That is, incentives for cooperation must be maintained at all times. If that was not the case then there would be a time after which only the stage-game Nash equilibrium would be played. Unravelling would then occur through backward induction and our equilibrium would have to consist of an infinite repetition of the stage-game Nash equilibrium.

Lemma 1. *Let (σ_1, σ_2) be a non-degenerate public Nash equilibrium of the dynamic game. Then for any $t \geq 0$, there exists $t' \geq t$ such that $\mathbb{P}_{\sigma}(a^{t'} \neq a^N) > 0$.*

Proof. Assume this is not the case and let $T \geq 0$ be the smallest integer such that for any $t' \geq T$ we have $\mathbb{P}_{\sigma}(a^{t'} \neq a^N) = 0$. Given that the public Nash equilibrium of the repeated game is non-degenerate it must be that $T > 0$. There is then a set of

[10] Recall that $\mathbb{P}_{\sigma}$ is the probability measure induced by the monitoring technology and a public strategy profile on the set of all public histories, as introduced at the end of Section 3.

[11] Note that because time is countable, this is equivalent to saying $\mathbb{P}_{\sigma}\{a^t = a^N, \forall t \geq 0\} = 1$.

[12] By non-degenerate Nash equilibrium, we mean an equilibrium in which stage-game payoffs are strictly higher in some periods than the payoff of the unique stage-game Nash equilibrium. That is, non-degenerate equilibria are equilibria that exhibit some degree of cooperation.

histories of positive measure such that an action profile different from a^N is played in period $T-1$ while a^N is played with probability one in the future. At least one player will then have an incentive to deviate in period $T-1$, a contradiction with the fact that (σ_1, σ_2) is an equilibrium. □

Note that there can still be histories after which the stage-game Nash equilibrium is played indefinitely. This set of histories however is of measure less than one.

Lemma 1 tells us that intertemporal incentives must remain throughout the game. However, as players are not mutually best-responding (since this only occurs at the unique stage-game Nash equilibrium), there will be myopic incentives to deviate. As discount factors become low those myopic incentives become more significant relative to intertemporal incentives. To mitigate this, action profiles have to approach the stage-game Nash equilibrium. The next lemma is the probabilistic version of Lemma 2.1 in BD.

Lemma 2. *Let* (σ_1, σ_2) *be a non-degenerate public Nash equilibrium of the dynamic game. Then the sequence of action profiles converges almost surely to the stage-game Nash equilibrium:* $\mathbb{P}_{\boldsymbol{\sigma}}(\lim_{t\to\infty} a^t = a^N) = 1$.

Proof. In any equilibrium, the instantaneous gain from a deviation must be lower than the maximal punishment incurred after a deviation: that is, for any $t \geq 0$ and $i = 1, 2$ we have that $0 \leq \max_{a_i'} u_i(a_i', a_j^t) - u_i(a^t) \leq \sum_{k=0}^{\infty} \beta_t^{t+1+k}(\overline{u}_i - \underline{u}_i)$ *a.s.*, where $\overline{u}_i = \max_{a\in A} u_i(a)$ and $\underline{u}_i = \min_{a\in A} u_i(a)$. As δ_t converges to zero, for any $\varepsilon > 0$ there exists T_ε such that for $t \geq T_\varepsilon$ we have $\beta_t^{t+1+k} \leq \varepsilon^{1+k}, k \geq 0$. Therefore $|\max_{a_i'} u_i(a_i', a_j^t) - u_i(a_i^t, a_j^t)| \to_{t\to\infty} 0$ *a.s.*, $i = 1, 2$. By the maximum theorem $a_j \mapsto \max_{a_i \in A_i} u_i(a_i, a_j)$ is continuous. As A is compact and there is a unique stage-game Nash equilibrium, this implies that $\mathbb{P}_{\boldsymbol{\sigma}}\{\lim_{t\to\infty} a^t = a^N\} = 1$. □

Corollary 1. *Let* (σ_1, σ_2) *be a non-degenerate public Nash equilibrium of the dynamic game. For any* $\varepsilon > 0$ *and* $\mu > 0$, $T_{\varepsilon,\mu} \geq 0$ *and a set* $\mathcal{H}_{\varepsilon,\mu}$ *of histories of length equal to* $T_{\varepsilon,\mu}$ *and of positive measure such that (i) for* $t \geq T_{\varepsilon,\mu}$, $\delta_t < \varepsilon$ *and (ii) for almost all histories* h *in* $\mathcal{H}_{\varepsilon,\mu}$ *and almost all histories* hh' *we have* $0 < \left\|\sigma(h) - a^N\right\| \leq \mu$ *and* $\left\|\sigma(hh') - a^N\right\| \leq \left\|\sigma(h) - a^N\right\|$.[13]

What Corollary 1 says is that we can find a set of histories of positive measure of a sufficiently long length and such that *(i)* the discount factor is arbitrarily close to zero; and *(ii)* all action profiles in this set are arbitrarily close but different from the stage-game Nash equilibrium and will not get further away in the future.

[13] Note that all norms in $\mathbb{R}^2$ are equivalent.

Finally, we will make use of the following lemma which gives us a lower bound on the instantaneous gains from a deviation when the action profile is close to the stage-game Nash equilibrium:

Lemma 3. *Under assumption 2, there exists* b, $\bar{\mu} > 0$ *such that for all* a *with* $\left\| a - a^N \right\| < \bar{\mu}$ *we have that*

$$\max_{x \in A_1} u_1(x, a_2) - u_1(a) + \max_{y \in A_2} u_2(a_1, y) - u_2(a) \geq b \left\| a - a^N \right\|^2$$

Lemma 3 tells us that as action profiles approach the stage-game Nash equilibrium, the sum of gain from each player best responding is at least of order two. The proof can be found in the appendix.

4.2. Proof of Theorem 2

Let us assume there is a non-degenerate public equilibrium (σ_1, σ_2) of the asymptotically finitely repeated game with imperfect public monitoring. We will show that at least one of the players must have a profitable deviation for histories in $\mathcal{H}_{\varepsilon,\mu}$. Recall that for histories in $\mathcal{H}_{\varepsilon,\mu}$, the discount factor is of order ε and action profiles are different from the stage-game Nash equilibrium but within a distance of μ'.

Note that at least one of the players must have myopic incentives to deviate, given that the action profile is different from the stage-game Nash equilibrium. Consider the following deviations for player i, $i = 1, 2$:

$$\sigma_i'(h^t) = \begin{cases} \phi_i(\sigma_j(h^t)) & \text{if } h^t \in \mathcal{H}_{\varepsilon,\mu}, \\ \sigma_i(h^t) & \text{otherwise.} \end{cases}$$

Strategy σ_i' prescribes myopically best responding to player j for any history in $\mathcal{H}_{\varepsilon,\mu}$, while agreeing with σ_i otherwise. (In particular, after having deviated in period t a player return to following the strategy prescribed by σ_i.) We now show that at least one of σ_1' or σ_2' must be a profitable deviation on $\mathcal{H}_{\varepsilon,\mu}$, that is that:

$$V_i(\sigma_i', \sigma_j \mid h^t) - V_i(\sigma_i, \sigma_j \mid h^t) > 0, \quad \forall h^t \in \mathcal{H}_{\varepsilon,\mu}, \tag{2}$$

for at least one of the players, where for $\sigma \in \{\sigma_i, \sigma_i'\}$, we can decompose $V_i(\sigma, \sigma_j \mid h^t)$ as follows:

$$V_i(\sigma,\sigma_j \mid h^t) = u_t(\sigma,\sigma_j \mid h^t) + \delta_t \int_Y V_i(\sigma_i,\sigma_j \mid h^t y)\mathbb{P}_Y(dy \mid \sigma,\sigma_j,h^t). \tag{3}$$

To do so we show that

$$\sum_{i=1}^{2}\{V_i(\sigma_i',\sigma_j \mid h^t) - V_i(\sigma_i,\sigma_j \mid h^t)\} > 0, \quad \forall h^t \in \mathcal{H}_{\varepsilon,\mu}, \tag{4}$$

which implies that (2) holds for at least one player.

Using (3), we can rewrite Equation (4) as follows, for an arbitrary $y_0 \in Y$:[14,15]

$$\begin{aligned}&\delta_t \sum_{i=1}^{2}\int_Y [V_i(\sigma_i,\sigma_j \mid h^t y) - V_i(\sigma_i,\sigma_j \mid h^t y_0)]\times \\ &\quad [\mathbb{P}_Y(dy \mid \sigma_i,\sigma_j,h^t) - \mathbb{P}_Y(dy \mid \sigma_i',\sigma_j,h^t)] \\ &\quad < \sum_{i=1}^{2}\{u_i(\sigma_i',\sigma_j \mid h^t) - u_i(\sigma_i,\sigma_j \mid h^t)\}.\end{aligned} \tag{5}$$

From Lemma 3 we know that for μ sufficiently small the right-hand side of (5) is bounded from below by $b\mu^2$, while each term in the sum on the left-hand side must be bounded from above by $c\mu^2$, for some constant $c > 0$ − this is because of Feller continuity (assumption 4), which bounds the difference between the probability measures by a term of order μ, and because all action profiles are within μ of a^N, which bounds the difference between the two value functions by another term of order μ (recall that the utility function are assumed to be twice-continuously differentiable − see Assumption 2). As δ_t is arbitrarily small (5) holds, which implies that at least one of the players has a profitable deviation on $\mathcal{H}_{\varepsilon,\mu}$.

V. Conclusion

In this paper we consider the class of repeated games with asymptotically finite horizons introduced by Bernheim and Dasgupta (1995) and show that non-degenerate equilibria are not robust to the introduction of an arbitrarily small amount of smooth public noise in the monitoring. This is because in any equilibrium of the perfect monitoring case, maximal punishments go to zero over

[14] This is because, for an arbitrary y_0, the term inside the integrals in what follows: $\int_Y V_i(\sigma_i, \sigma_j \mid h^t y_0)\mathbb{P}_Y(dy \mid \sigma_i,\sigma_j,h^t) - \int_Y V_i(\sigma_i,\sigma_j \mid h^t y_0)\mathbb{P}_Y(dy \mid \sigma_i',\sigma_j,h^t) = V_i(\sigma_i,\sigma_j \mid h^t y_0) - V_i(\sigma_i,\sigma_j \mid h^t y_0) = 0$.

[15] For ease of notation, we write $u_i(\sigma_i,\sigma_j \mid h^t)$ instead of $u_i(\sigma_i(h^t),\sigma_j(h^t))$. Similarly, we use $\mathbb{P}_Y(dy \mid \sigma_i,\sigma_j,h^t)$ to denote the distribution $\mathbb{P}_Y(dy \mid \sigma_i(h^t),\sigma_j(h^t))$.

time, as discount factors go to zero. The introduction of a small amount of noise in the monitoring renders deviations profitable.

Even though the rate at which punishments go to zero is exogenous, given by the rate of decline of discount factors, it is not obvious that the result extends to the case of private strategies or private monitoring. Consider for example Lemma 1. Under private monitoring, even though players do not play the unique stage-game Nash equilibrium, they could still both be best-responding to their beliefs about the other player's strategy.

A. Proof of Lemma 3

Let $u_i^d(a_i, a_j): A \to \mathbb{R}$ be the function which returns the highest payoff player i can get by deviating from action profile (a_i, a_j), that is $u_i^d(a_i, a_j) = u_i(\phi_i(a_j), a_j)$. Consider the following second order Taylor expansions around a^N:

$$\begin{aligned}
u_1(a_1, a_2) &= u_1(a^N) + (a_2 - a_2^N)\frac{\partial u_1}{\partial a_2}(a^N) + (a_1 - a_1^N)^2 \frac{1}{2}\frac{\partial^2 u_1}{\partial a_1^2}(a^N) \\
&\quad + (a_2 - a_2^N)^2 \frac{1}{2}\frac{\partial^2 u_1}{\partial a_2^2}(a^N) + (a_1 - a_1^N)(a_2 - a^N)\frac{\partial^2 u_1}{\partial a_1 \partial a_2}(a^N) \\
&\quad + o(\|a - a^N\|^2),
\end{aligned}$$

where the first order term in a1 is zero from the first order condition $\frac{\partial u_1}{\partial a_1}(a^N) = 0$ and

$$\begin{aligned}
u_1^d(a_1, a_2) &= u_1(\phi_1(a_2), a_2) \\
&= u_1(\phi_1(a_2^N), a_2^N) + (a_2 - a_2^N)\left[\frac{\partial u_1}{\partial a_1}(\phi_1(a_2^N), a_2^N)\phi_1'(a_2^N) + \frac{\partial u_1}{\partial a_2}(\phi_1(a_2^N), a_2^N)\right] \\
&\quad + (a_2 - a_2^N)^2 \frac{1}{2}\left[\frac{\partial^2 u_1}{\partial a_1 \partial a_2}(\phi_1(a_2^N), a_2^N)\phi_1'(a_2^N) + \frac{\partial^2 u_1}{\partial a_2^2}(\phi_1(a_2^N), a_2^N)\right] \\
&\quad + o(\|a - a^N\|^2) \\
&= u_1(a^N) + (a_2 - a_2^N)\frac{\partial u_1}{\partial a_2}(a^N) + \\
&\quad (a_2 - a_2^N)^2 \frac{1}{2}\left[\frac{\partial^2 u_1}{\partial a_1 \partial a_2}(a^N)\phi_1'(a_2^N) + \frac{\partial^2 u_1}{\partial a_2^2}(a^N)\right] \\
&\quad + o(\|a - a^N\|^2),
\end{aligned}$$

where the last equality is obtained because from the envelope theorem we have that $\frac{\partial u_1}{\partial a_1}(\phi(a_2^N), a_2^N) = 0$.

Therefore

$$\begin{aligned}
u_1^d(a_1, a_2) - u_1(a_1, a_2) &= (a_2 - a_2^N)^2 \frac{1}{2}\frac{\partial^2 u_1}{\partial a_1 \partial a_2}(a^N)\phi_1'(a_2^N) \\
&\quad - (a_1 - a_1^N)^2 \frac{1}{2}\frac{\partial^2 u_1}{\partial a_1^2}(a^N) - (a_1 - a_1^N)(a_2 - a_2^N)\frac{\partial^2 u_1}{\partial a_1 \partial a_2}(a^N) + o(\|a - a^N\|^2),
\end{aligned}$$

which can be further simplified into

$$u_1^d(a_1,a_2)-u_1(a_1,a_2)=-\frac{1}{2}\frac{\partial^2 u_1}{\partial a_1^2}(a^N)[(a_2-a_2^N)^2(\phi_1'(a_2^N))^2$$
$$+(a_1-a_1^N)^2-2(a_1-a_1^N)(a_2-a_2^N)\phi_1'(a_2^N)]+o(\|a-a^N\|^2),$$

by noting that $\frac{\partial^2 u_1}{\partial a_i \partial a_j}=-\phi_i'(a_2^N)\frac{\partial^2 u_i}{\partial a_i^2}$ from differentiating the first order condition.

We obtain a similar expression for $u_2^d(a_1,a_2)-u_2(a_1,a_2)$ and sum both gains to obtain:

$$u_1^d(a)-u_1(a)+u_2^d(a)-u_2(a)=$$
$$\frac{1}{2}(a-a^N)^t[D\phi(a^N)-I]^t U[D\phi(a^N)-I](a-a^N)+o(\|a-a^N\|^2),$$

where

$$D\phi(a^N)-I=\begin{pmatrix} -1 & \phi_1'(a_2^N) \\ \phi_2'(a_1^N) & -1 \end{pmatrix},$$

and

$$U=-\begin{pmatrix} \frac{\partial^2 u_1}{\partial a_1^2}(a^N) & 0 \\ 0 & \frac{\partial^2 u_2}{\partial a_2^2}(a^N) \end{pmatrix}.$$

The matrix $[D\phi(a^N)-I]^t U[D\phi(a^N)-I]$ is positive definite as we know that when $a \neq a^N$ then at least one player has a profitable deviation and therefore $u_1^d(a)-u_1(a)+u_2^d(a)-u_2(a)>0$.[16] Moreover its determinant is non-zero as $D\phi(a^N)-I$ and U are non-singular. U is non-singular as utility functions are strictly concave at the Nash equilibrium, so that $\frac{\partial^2 u_i}{\partial a_i^2}(a^N)\neq 0$, $i=1,2$. Let $\lambda_{\min}$ denote the smallest eigenvalue of $[D\phi(a^N)-I]^t U[D\phi(a^N)-I]$, which is strictly positive. We then have the following inequality:

$$u_1^d(a)-u_1(a)+u_2^d(a)-u_2(a)\geq\frac{1}{2}\lambda_{\min}\|a-a^N\|^2+o(\|a-a^N\|^2). \tag{6}$$

[16] The matrix $[D\phi(a^N)-I]^t U[D\phi(a^N)-I]$ cannot be zero as Assumption 2 implies it must have full rank.

Assume now that the conclusion of Lemma 3 does not hold. This implies that for all b, $\bar{\mu} > 0$, there is an a with $\left\| a - a^N \right\| < \bar{\mu}$ and such that we have that $u_1^d(a) - u_1(a) + u_2^d(a) - u_2(a) < b \left\| a - a^N \right\|^2$, which would contradict (6).

References

Bernheim, B. and A. Dasgupta (1995), “Repeated Games with Asymptotically Finite Horizons,” *Journal of Economic Theory*, 67(1), 129–152.

Guéron, Y. (2015), “Failure of Gradualsim under Imperfect Monitoring,” *Journal of Economic Theory*, 157, 128–145.