Folk Theorems[1]

Olivier Compte and Andrew Postlewaite

First version: January 2013. This version: April 2013

Abstract:

Much of the repeated game literature is concerned with proving Folk Theorems. The logic of the exercise is to specify a particular game, and to explore for that game specification whether any given feasible (and individually rational) value vector can be an equilibrium outcome for some strategies when agents are sufficiently patient. A game specification includes a description of what agents observe at each stage. This is done by defining a monitoring structure, that is, a collection of probability distributions over the signals players receive (one distribution for each action profile players may play). Although this is simply meant to capture the fact that players don't directly observe the actions chosen by others, constructed equilibria often depend on players precisely knowing these distributions, somewhat unrealistic in most problems of interest. We revisit the classic Folk Theorem for games with imperfect public monitoring, asking that incentive conditions hold not only for a precisely defined monitoring structure, but also for a ball of monitoring structures containing it. We show that efficiency and incentives are no longer compatible.

1 Introduction

The repeated game literature studies long run/repeated interactions, aiming to understand how repetition may foster cooperation. Conditioning behavior on observations is an essential ingredient, and the literature has tried to understand how the nature and quality of observations affect cooperation possibilities. Specifically, the analysis starts with a stage game characterized by a payoff structure describing how action profiles affect gains for each

[1]We thank Drew Fudenberg, George Mailath, Steve Morris, Larry Samuelson and Yuichi Yamamoto for helpful comments.

player, and a monitoring structure, that is, a probability distribution over signals for each action profile possibly played that captures the possibility that the actions played are not perfectly observable. Taking payoff and monitoring structures as given, one then attempts to characterize the set of (sequential) equilibria of the repeated game, aiming for a Folk theorem.

Although an imperfect monitoring structure is just a modelling device employed to capture an agent's inability to observe perfectly what others are doing, equilibrium constructions sometimes hinge on the precise specification of that monitoring structure, with strategies finely tuned to that particular specification, as though agents could easily determine the precise (stochastic) relationship between the action profile played and the signals observed, despite the fact that others' actions are not observable. This seems unrealistic, more so when the environment that agents face varies over time. Besides, one suspects that the plethora of equilibria one can construct might be a consequence of this presumed unlimited ability of agents to tailor their strategies to the underlying parameters of the game.

In a companion paper (Compte and Postlewaite (2013)), we have illustrated the lack of robustness of some of these equilibrium constructions (belief free equilibria), by considering an environment in which there are exogenous and persistent shocks to the monitoring structure. In this paper, we apply a similar methodology to repeated games in which the signals that players receive are public, with a similar motivation: agents cannot plausibly know with precision the underlying monitoring technology, and the range of possible stochastic processes over monitoring technologies is so vast that one cannot plausibly assume agents can learn precisely, nor track variations in, the underlying monitoring technology.[2]

Specifically, we shall restrict attention to equilibria that are robust to a *rich* set of monitoring technologies. By this we mean equilibria in which players need not try to make inferences about the actual monitoring technology within that set.[3] We emphasize that we do not think of an arbitrarily

[2] In this respect, we differ from the perspective adopted in Fudenberg and Yamamoto (2010). Fudenberg and Yamamoto start with a concern similar to ours, yet they propose a set up in which, over time, players manage to learn the underlying monitoring technology.

[3] Stronger restrictions would obtain if one considered variations over time in the moni-

small set of monitoring technologies, because we have in mind that agents do not have a precise (nor commonly shared) idea of the monitoring structure they face.

Our main insight concerns the equilibrium constructions proposed by Fudenberg, Levine and Maskin (1994) (FLM hereafter). In FLM, Folk theorems are obtained by constructing continuation payoff vectors lying on one side of a hyperplane but arbitrarily close to the hyperplane. Even if one can construct a strategy with appropriate continuation values for a given monitoring structure, there is no guarantee that continuation values will continue lying on the appropriate side of (nor close to) the hyperplane when the monitoring structure varies.

We illustrate the difficulty in two ways: by proving an impossibility result (Section 2) and by constructing a numerical example in which we vary the monitoring technology (Section 3), showing that while some equilibrium constructions have good efficiency properties for a given monitoring technology, they lose much of their appeal when the monitoring structure varies.

2 The model

Consider a repeated game between n players, each player i simultaneously choosing $a_i \in A_i$ in each stage game. We denote by g^a the expected gains associated with action profile $a = (a_i)_i$. We assume there is a unique action profile a^* maximizing the sum of players' payoffs. That is, there exists $D > 0$ such that, for any other action profile a, we have:

$$g^* \equiv \sum_i g_i^{a^*} \geq \sum_i g_i^a + D.$$

We also assume that a^* is not a Nash equilibrium of the stage game. Consequently, some player i has a profitable deviation $\widehat{a}_i$ that yields a *strictly*

toring technology, and if one insisted on robustness with respect to the stochastic process that determines the monitoring technology. But negative results already obtain under the weaker requirement above that implicitly restrict attention to stochastic processes in which the monitoring structure does not change over time.

positive gain, which we denote d_i:

$$d_i \equiv \max_{a_i} g_i^{a_i, a^*_{-i}} - g_i^{a^*}.$$

After choices have been made, players receive a public signal y drawn from a finite set Y.[4] An *imperfect public monitoring structure* q specifies for each a a distribution over public signals: $q = (q_a)_{a \in A}$ with each $q_a \in \Delta(Y)$. We consider a *rich* set of monitoring structures, denoted Q, each $q \in Q$ having *full support.* By full support we mean that there exists M such that:

$$\max_{a,y,y'} \frac{q_a(y)}{q_a(y')} \leq M.$$

By *rich*, we mean that Q contains a ball $B \subset [\Delta(Y)]^A$, centered on some q^0. Specifically, for any $m > 0$, we define

$$B^m = \{q \in [\Delta(Y)]^A, \text{ for all } a, y, \frac{q_a(y)}{q_a^0(y)} \in [1-m, 1+m]\}$$

and assume that $Q \supset B^m$ for some $m > 0$. We shall refer to m as the size of the ball B^m.

It is convenient to refer to $\theta \in \Theta$ as the "current" state of the monitoring structure, as in principle there could be variations *over time* in the monitoring technology. For the purpose of this paper however, it will not be necessary to allow for such variations over time: we will thus identify Θ with Q, and refer to either θ or q^θ as the underlying monitoring technology.

The game is repeated, and players evaluate payoffs using the same discount factor $\delta < 1$. We consider pure public strategies.[5] We denote by $v_i^\theta(\sigma)$ the value that player i obtains under monitoring structure θ when players follow the strategy profile σ, and we define $\mathcal{E}^{\delta,\theta}$ as the set of perfect pure public equilibria when the monitoring structure is q^θ. For reasons explained in the introduction, we are interested in the set of $\Theta-$robust equilibria, i.e.

[4] The assumption that Y is finite is not essential however.

[5] That is, strategies in which play is only conditioned on the history of past public signals, and in which play is in pure strategies. The proof actually only requires that play be in pure strategies a fraction of the time bounded away from 0. See footnote 13.

$\mathcal{E}^{\delta} = \cap_{\theta \in \Theta} \mathcal{E}^{\delta,\theta}$.[6]

Before continuing, note that we keep the payoff structure unchanged across all monitoring structures $q \in Q$. This assumption is, for example, consistent with players getting private signals about own payoffs, and with the public signal being a noisy or crude aggregate of the private signal profile received by players.[7] This assumption however is not essential to our argument.[8]

Having fixed the parameters m, M, d_i and D, our main proposition states that inefficiencies must arise:

> **Proposition 1**: *There exists $\eta_0 > 0$ such that for any δ and any $\sigma \in \mathcal{E}^{\delta}$, $\sum_i v_i^{\theta}(\sigma) < \sum_i g_i^{a^*} - \eta_0$ for at least one θ.*

In the Appendix, we prove a stronger result, showing that the statement above holds for a set of monitoring structures of positive measure. One can also show that for a ball of small size m, the inefficiency bound η_0 may be chosen linear in m.

Before presenting the argument, we introduce some notation. For any given i and $\alpha_i \in \Delta(A_i)$, we define the function ϕ_{i,α_i} that maps any monitoring structure q to q' so that:

$$\begin{aligned} q'_a &= q_a \text{ for } a \neq a^* \text{ and} \\ q'_{a^*} &= q_{\alpha_i, a^*_{-i}}. \end{aligned}$$

[6] Thus, the path proposed is in the spirit of the ex post notion of equilibrium that Fudenberg and Yamamoto (2010) consider. This path differs from standard approaches to robustness (such as those pursued in Chassang and Takahashi (2011), following Kajii and Morris (1997)), in two ways. (i) We do not have in mind arbitrarily small perturbations in the monitoring technology. (ii) We are not asking whether *nearby strategy profiles* support the same value vector when the monitoring technology varies. We analyze the consequence of varying the monitoring technology, *for a given strategy profile.*

[7] Formally, assume each player i receives a private signal z_i about his payoff, according to some joint distribution $h_a(z)$ over signal profiles $z = (z_i)_i$, and that in addition, players receive a public signal that consists of a coarse/noisy aggregate of the private signal profile $z = (z_i)_i$, drawn from the distributions $k_z^{\theta}(y)$ for each z. This generates a payoff structure $g_i^a = E_{h^a} r_i(a_i, z_i)$, and a public monitoring structure $q_a^{\theta} \equiv \sum_z k_z^{\theta}(y) h_a(z)$.

[8] See the Appendix where we allow the payoff structure to depend on the monitoring structure.

Next we define the set:

$$\Gamma_i = \{\alpha_i \in \Delta(A_i) \mid \phi_{i,\alpha_i}(q^0) \in Q\}.$$

The set Γ_i characterizes the degree to which player i's deviations from a^* can go unnoticed, in the sense that they would generate the same distribution over outcomes under q^0 as a^* would under some other monitoring structure q in Q. Γ_i is not empty because it contains a_i^*,[9] and since Q contains B^m, Γ_i contains all mixtures (sufficiently) near a_i^*.[10]

Finally, we let:

$$\Delta_i = \max_{\alpha_i \in \Gamma_i} g_i(\alpha_i, a_{-i}^*) - g_i(a^*).$$

Since Γ_i contains all mixtures near a_i^*, and since a^* is not a Nash equilibrium, Δ_i is strictly positive for some player i.[11]

The argument for Proposition 1 builds on Δ_i being positive. Intuitively, it runs as follows. If an equilibrium profile σ calls for efficient behavior a^* most of the time under both q^0 and $q^1 = \phi_{i,\alpha_i}(q^0)$, then it means that under q^0, it asks for a^* most of the time, whether player i plays a_i^* or α_i when the prescribed play is a^*. Since $\Delta_i > 0$, playing α_i rather than a_i^* is strictly preferred by player i.

Proof of Proposition 1: For any η, δ and $\sigma \in \mathcal{E}^\delta$, define $Q^{\eta,\delta,\sigma}$ as the set of monitoring structures θ for which $\sum_i v_i^\theta(\sigma) > \sum_i g_i^{a^*} - \eta$. We first observe that for any $\theta \in Q^{\eta,\delta,\sigma}$, the path induced by σ must call for playing a^* most of the time, that is except for a fraction of the time equal to η/D.[12]

[9] Note that for $\alpha_i = a_i^*$, ϕ_{i,α_i} is the identity function.

[10] It is sufficient that some weight $1-\alpha_i$ with $\alpha_i \leq \frac{m}{M-1}$ be put on a_i^*. Indeed, $q^0_{\alpha_i,a_{-i}^*} \equiv \alpha_i q^0_{\hat{a}_i,a_{-i}^*} + (1-\alpha_i)q^0_{a^*}$, so $\frac{q_{a^*}}{q^0_{a^*}} = \frac{q^0_{\alpha_i,a^*_{-i}}}{q^0_{a^*}} = 1-\alpha_i + \alpha_i \frac{q^0_{\hat{a}_i,a^*_{-i}}}{q^0_{a^*}} \in [1-\alpha_i(1-1/M), 1+\alpha_i(M-1)]$, hence $\frac{q_{a^*}}{q^0_{a^*}} \in [1-m, 1+m]$ when $\alpha_i \leq \frac{m}{M-1}$.

[11] From footnote 10, one actually derives $\Delta_i \geq d_i m/(M-1)$.

[12] This is because $v^\theta(\sigma)$ can be written as a weighted average of the action profiles played in equilibrium: $v^\theta(\sigma) = \sum \lambda^\theta_{\sigma,a} g^a$, with $\lambda^\theta_{\sigma,a}$ appropriately taking into account the date at which a is played and the discount factor. We have $\sum v_i^\theta(\sigma) \leq \sum_i g_i^{a^*} - (1-\lambda^\theta_{\sigma,a^*})D$, implying that $\lambda^\theta_{\sigma,a^*} \geq 1-\eta/D$.

Let $\overline{g}$ and $\underline{g}$ denote upper and lower bounds on players' payoffs. In what follows, we set

$$\eta_0 = \Delta_i \frac{D}{4(\overline{g} - \underline{g})}$$

and we assume by contradiction that $Q \subset Q^{\eta_0,\delta,\sigma}$.

Consider the strategy $\sigma_i^{\alpha_i}$ for player i that plays α_i whenever players are supposed to choose a^*. Consider the mixture $\alpha_i \in \Gamma_i$ that achieves the gain Δ_i, and the monitoring structure $\theta = \phi_{i,\alpha_i}(q^0)$. By definition of Γ_i, $\theta \in Q$ and the distribution over signals induced by (α_i, a^*_{-i}) under q^0 coincides with that induced by (a^*_i, a^*_{-i}) under θ. So the two paths induced by $(\sigma_i^{\alpha_i}, \sigma_{-i})$ under q^0 and σ under θ coincide except that whenever a^* is played under θ, (α_i, a^*_{-i}) is played under q^0. Since $\theta \in Q \subset Q^{\eta_0,\delta,\sigma}$, the path induced by σ under θ (hence both paths) must prescribe that players choose a^* a fraction $1 - \eta_0/D$ of the time. We thus have:[13]

$$v_i^{q^0}(\sigma_i^{\alpha_i}, \sigma_{-i}) \geq (g_i^{a^*} + \Delta_i)(1 - \eta_0/D) + \eta_0 \underline{g}/D \geq g_i^{a^*} + \Delta_i - \eta_0(\overline{g} - \underline{g})/D \geq g_i^{a^*} + 3\Delta_i/4.$$

Since $\sigma \in \mathcal{E}^{\delta,q^0}$, we have:

$$v_i^{q^0}(\sigma_i^{\alpha_i}, \sigma_{-i}) \leq v_i^{q^0}(\sigma)$$

thus implying:

$$v_i^{q^0}(\sigma) \geq g_i^{a^*} + 3\Delta_i/4.$$

To conclude, observe that since $q^0 \in Q \subset Q^{\eta_0,\delta,\sigma}$, the path induced by σ under q^0 prescribes that players choose a^* a fraction $1 - \eta_0/D$ of the time, hence:

$$v_i^{q^0}(\sigma) \leq g_i^{a^*}(1 - \eta_0/D) + \overline{g}\eta_0/D \leq g_i^{a^*} + \Delta_i/4$$

contradicting the previous inequality. QED

[13] If in equilibrium, the action profile a^* were played in a pure strategy only a fraction of the time β, we would consider the deviation $\sigma_i^{\alpha_i}$ that plays α_i only in these periods, hence we would then replace Δ_i with $\beta\Delta_i$, and the argument we propose would still be valid.

3 An illustration

We next present an example that provides intuition as to why the typical FLM construction fails to be robust in our sense. In FLM, the equilibrium strategies that achieve an outcome close to some efficient outcome must be constructed so that continuation values follow a random walk (with no drift) on the frontier of some smooth self generating set[14] lying close to the point on the efficient frontier one wants to sustain. When one modifies the monitoring structure, that same strategy profile may be robust in the sense that it still defines an equilibrium. However, the strategy profile now generates a process over continuation values that *drifts*, so that even if incentives are preserved, the location of the self generating set no longer lies close to the original one.

We illustrate this phenomenon with a simple two player example. We consider a stage game with a symmetric payoff structure, an action profile a^* yielding the maximum feasible joint payoff (g, g), and two distinct *Nash profiles* b^1 and b^2 yielding payoffs $(0, g + \gamma)$ and $(g + \gamma, 0)$ respectively.[15]

We wish to support equilibrium payoffs close to (g, g). There are two signals available y_1 and y_2. We let

$$p = \Pr\{y_2 \mid a^*\}$$

and we assume that any deviation from a^* by player 1 increases the likelihood of y_1, while any deviation from a^* by player 2 increases the likelihood of y_2. Formally, we assume that:

$$\Pr\{y_i \mid a_i, a^*_{-i}\} > q > \max(p, 1 - p).$$

In other words, the signals permit one to distinguish statistically between player 1 and 2's deviations. Our objective is to design equilibria that permit

[14] Self generating sets have been introduced by Abreu, Pearce and Stachetti (1986,1990). A set W is self-generating if all values in W can be sustained and enforced with continuation values also lying in W.

[15] The assumption that there are two asymmetric Nash equilibria is for convenience, so that we need not worry about supporting these asymmetric continuation values. In general, such asymmetric points have to be supported, and the difficulties we report for sustaining values close to (g, g) would also arise for these asymmetric points.

one to approach (g, g) when $p = 1/2$, and then to examine the effect of variations in the monitoring technology.

We shall construct equilibria in which only a^*, b^1 and b^2 are played, and in which occasionally either b^1 or b^2 is triggered for a while, based on the balance between the number of occurrences of y_1 and y_2.

Specifically, we describe the strategies we consider as follows. Each player starts with K credits. Each signal y_i generates a transfer of one credit from player i to player j. When a player, say player i, has no more credits, then b^i is played. At the end of that period, player i gets a transfer from player j of one unit of credit with probability μ.[16]

The strategy considered is thus characterized by the pair (K, μ), with K affecting the length of the cooperative phase, and μ affecting the length of the punishment phase.

Continuation values depend only on the balance of credits. Letting k_i denote the number of credits for player i, we define $n = (k_1 - k_2)/2 \in \{-K, .., 0, ...K\}$ and denote by v^n the corresponding continuation value. For all $n \neq K, -K$ we have:

$$v_1^n = (1-\delta)g + \delta(pv_1^{n-1} + (1-p)v_1^{n+1})$$

and

$$v_1^{-K} = \delta(\mu v_1^{-K+1} + (1-\mu)v_1^{-K}) \text{ and } v_1^K = (1-\delta)\gamma + \delta(\mu v_1^{K-1} + (1-\mu)v_1^K).$$

We turn to specific numerical values to illustrate our point. We fix $\delta = 0.99$, $g = 3$ and $\gamma = 0.5$, and consider the strategy defined by $K = 10$ and $\mu = 0.95$. The following graph depicts continuation values for various specification of the monitoring technology: $p = 1/2, 0.45, 0.4$.[17]

[16] This assumes a public randomization device. One could easily dispense of that assumption however, for example assuming that player i gets a unit of credit after b^i has been played for κ periods.

[17] Each black (respectively red and blue) dot corresponds to the locus of a particular continuation value vector $v^n = (v_1^n, v_2^n)$ when $p = 1/2$ (respectively 0.45 and 0.4). The size of the dot is proportional to the long run probability n.

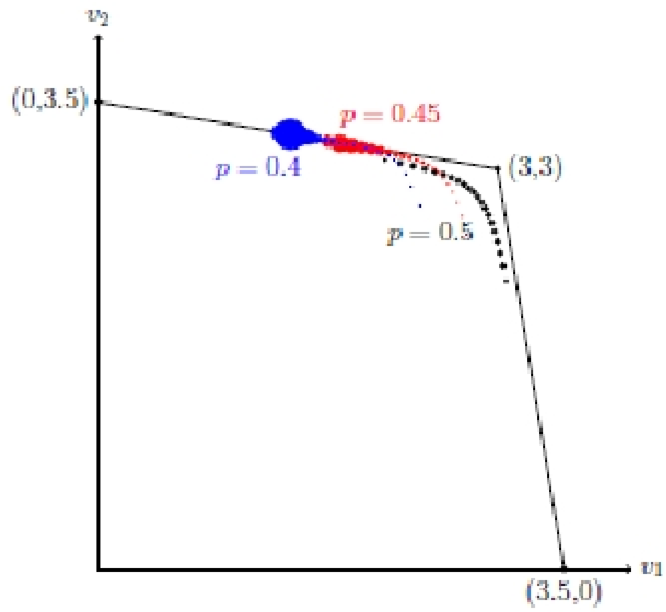

Figure 1: Location of W for $p = 0.5$, $p = 0.45$ and $p = 0.4$.

When $p = 1/2$, punishment phases are triggered on somewhat rare occasions (because K is relatively large), and a payoff close to the efficient outcome $g^{a^*} = (3, 3)$ can be sustained. The set of value vectors $W = \{v^{-K}, .., v^0, v^K\}$ is a self generating set,[18] and over time, continuation values follow a random walk on this set.[19]

When $p \neq 1/2$, say $p < 1/2$, the proposed strategies may still remain in equilibrium, that is, the induced set of continuation values $W^p =$

[18] The differences $v_i^{n+1} - v_i^{n-1}$ are at least equal to 0.02. So enforcement in ensured if q is large enough and the gain from deviating from a^* small enough.

[19] Note that in the long run, all $n \neq K, -K$ are equally likely.

Also note that one may adjust the provision of incentives by varying μ. A smaller μ means a smaller chance to move away from the punishment phase. It would imply that all values shift downward, and it would also imply larger differences $v_n - v_{n-1}$, thus stronger incentives.

$\{v^{-K}, .., v^0, v^K\}$ may still be a self generating set. However credit transfers are no longer adapted to that monitoring technology. As a result, the number of credits player 1 has drifts toward lower values (because lower p means more frequent signals y_1), the punishment phase (b^1) is often triggered, and the location of W^p moves away from (g, g). When $p = 0.45$ or $p = 0.4$, the punishment phase (b^2) is almost never triggered, and equilibrium play essentially consists of a combination of b^1 and a^*. The insight of Radner, Myerson and Maskin (1986) applies, with equilibrium values necessarily bounded away from $g(a^*)$.

A similar analysis holds for $p > 1/2$, with continuation values drifting towards the punishment (b^2).

4 Conclusion

The usual methodology assumes a precise monitoring technology, and investigates how players can best take advantage of that monitoring technology, finding strategies that preserve incentives and yet seldom induce inefficient punishment phases. For example, the literature has been helpful in understanding how the inefficiencies identified by Radner, Myerson and Maskin (1986) could be alleviated when players have access to a sufficiently rich monitoring structure that permits statistically distinguishing between each player's deviations (Fudenberg, Levine and Maskin (1994), or when signals are only revealed with delay (Abreu, Milgrom and Pearce (1991).

While these insights are important, the quest for Folk theorems sometimes pushes the analyst toward fine tuning strategies precisely to the particular monitoring structure that agents face: the actions are not observable, yet the strategies may be fine tuned to the distributions that generate signals, as though these distributions were perfectly observed. When such fine tuning cannot be done, for example because the environment that players face varies or because information about the monitoring technology and about changes in the monitoring technology is limited, difficulties arise. In a companion paper, we have shown that belief free constructions in the literature fail to be robust in our sense. In this paper, it is not the equilibrium construction itself that fails to be robust. Rather, it is the ability to sustain a given Pareto efficient point that fails to be robust; because equilibrium strategies cannot be tailored to each different realization of the underlying state (i.e. the monitoring structure), the locus of each state contingent equilibrium values drifts away from the point that one would have liked to sustain.

References

Abreu, D., P. Milgrom and Pearce, D. (1991) "Information and Timing in Repeated Partnerships", *Econometrica* , 59, pp. 1713-1733

Abreu, D., D. Pearce, and E. Stacchetti (1986) "Optimal Cartel Equilibria With Imperfect Monitoring," *Journal of Economic Theory*, 39, 251–269.

——— (1990) "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041-1063.

Chassang, S. and S. Takahashi, S. (2011), "Robustness to incomplete information in repeated games," *Theoretical Economics*, 6: 49–93.

Compte, O. and A. Postlewaite (2013), "Belief free equilibria", mimeo

Fudenberg, D., D. K. Levine, and E. Maskin (1994) "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62, 997-1040.

Fudenberg, D., and Y. Yamamoto (2010) "Repeated Games where the payoffs and monitoring structure are unknown", *Econometrica* 78, pp. 1673-1710.

Kajii, A. and S. Morris (1997), "The robustness of equilibria to incomplete information." *Econometrica*, 65, 1283–1309.

Radner, R., R. Myerson, and E. Maskin (1986) "An Example of a Repeated Partnership Game with Discounting and with Uniformly Inefficient Equilibria", *Review of Economic Studies* 53, 59-69.

Appendix

We report here a stronger result. In essence, it builds on the observation that *there are actually many monitoring structures* for which player i's deviations from a^* can go unnoticed. We also present a proof in which the payoff structure is allowed to depend on the monitoring structure θ.

Specifically, we define $g_i^\theta(a)$ to make explicit that payoffs may depend on the monitoring structure, and assume that for some player i and action $\widehat{a}_i$, $g_i^\theta(\widehat{a}_i, a_{-i}^*) > g_i^\theta(a^*)$ for all $\theta \in Q$, that is, $\widehat{a}_i$ is a profitable deviation for all monitoring structures in Q.

> **Proposition 2**: *There exists $d, \eta_0 > 0$ such that for any δ and any $\sigma \in \mathcal{E}^\delta$, $\sum_i v_i^\theta(\sigma) < \sum_i g_i^\theta(a^*) - \eta_0$ for a set of monitoring structures of measure at least equal to d.*

Proof: Define B^m to be the ball of monitoring structures around q^0 of size m.Without loss of generality, we consider the case where $Q = B^{2m}$ for some $m > 0$. For simplicity, we shall also abuse notation and refer to generic monitoring structures as either θ or q.

We start with some new notation. For any $\alpha_i \in [0, 1]$, we define (abusing notation) α_i as the mixed action for player i that puts weight α_i on the profitable deviation $\widehat{a}_i$ and weight $1 - \alpha_i$ on a_i^*.

Next, for any given i and α_i, we define the function ϕ_{i,α_i} that maps any monitoring structure q to q' so that:

$$\begin{aligned} q'_a &= q_a \text{ for } a \neq a^* \text{ and} \\ q'_{a^*} &= q_{\alpha_i, a_{-i}^*} \equiv \alpha_i q_{\widehat{a}_i, a_{-i}^*} + (1 - \alpha_i) q_{a^*}. \end{aligned}$$

Note that for $\alpha_i = 0$, ϕ_{i,α_i} is the identity function, and that, since distributions have positive support, the function ϕ_{i,α_i} can be inverted if α_i is sufficiently small.[20] In what follows, we fix $\alpha_i^m > 0$ so that $\phi_{i,\alpha_i^m}^{-1}(B^m) \subset B^{2m} (= Q)$ and set $\Delta_i^m = \min_{\theta \in B^{2m}} g_i^\theta(\alpha_i^m, a_{-i}^*) - g_i^\theta(a^*)$.

For any η, δ and any $\sigma \in \mathcal{E}^\delta$, we now define $Q^{\sigma,\eta}$ as the (convex) set of monitoring structures $\theta \in B^m$ for which $\sum_i v_i^\theta(\sigma) > \sum_i g_i^\theta(a^*) - \eta$. Note

[20] This is because starting from $q' \in B^m$, one may define $q_{a^*} = (q'_{a^*} - \alpha_i q'_{\widehat{a}_i, a_{-i}^*})/(1 - \alpha_i)$. Since all $q \in Q$ have full support on Y, there exists $\gamma > 0$ such that $q'(y) \geq \gamma$ for all $q' \in B^m$ and $y \in Y$. So for α_i small enough ϕ_{i,α_i}^{-1} is well defined on B^m.

that for any such θ, σ must call for playing a^* most of the time (on the equilibrium path), that is, except for a fraction of the time of order η.[21]

We now choose η_0 small compared to Δ_i^m. Our aim is to show that for any δ and $\sigma \in \mathcal{E}^\delta$, the set $\widehat{Q}$ of monitoring structures $\theta \in Q$ for which $\sum_i v_i^\theta(\sigma) \leq \sum_i g_i^{\theta,a^*} - \eta_0$ has measure bounded away from 0. Our main observation is as in the proof of Proposition 1:

$$\phi^{-1}_{i,\alpha_i^m}(Q^{\sigma,\eta_0}) \cap Q^{\sigma,\eta_0} = \varnothing. \tag{1}$$

Once this is proved, the conclusion obtains, as we now show. First observe that (1) implies that $\widehat{Q}$ contains $\phi^{-1}_{i,\alpha_i^m}(Q^{\sigma,\eta_0})$ and the complement of Q^{σ,η_0} in B^m, [22] that is:

$$\widehat{Q} \supset \phi^{-1}_{i,\alpha_i^m}(Q^{\sigma,\eta_0}) \cup (B^m \backslash Q^{\sigma,\eta_0}).$$

Next we distinguish two cases. Either Q^{σ,η_0} has a small measure relative to B^m (i.e. smaller than a fraction $1-d$ with d small), and the desired conclusion obtains: $\mu(\widehat{Q}) \geq d\mu(B^m)$; or Q^{σ,η_0} has a large measure relative to B^m (larger than a fraction $1-d$). For d small enough (compared to m), any convex set having a measure larger $1-d$ relative to B^m must contain the smaller ball $B^{m/2}$. Since Q^{σ,η_0} is a convex set, we obtain $\phi^{-1}_{i,\alpha_i^m}(Q^{\sigma,\eta_0}) \supset \phi^{-1}_{i,\alpha_i^m}(B^{m/2})$. So $\phi^{-1}_{i,\alpha_i^m}(Q^{\sigma,\eta_0})$ has a measure bounded away from 0, and the conclusion obtains in that case too.

The proof of (1) follows the steps of the proof of Proposition 1. Assume by contradiction that there exists θ and θ' both in Q^{σ,η_0} such that $\theta' = \phi^{-1}_{i,\alpha_i^m}(\theta)$. Then by construction, the distribution over signals induced by (α_i, a^*_{-i}) under θ coincides with that induced by (a_i^*, a^*_{-i}) under θ'. Now define the strategy σ_i^m for player i that plays α_i^m whenever players are supposed to choose a^*. The two paths induced respectively by $(\sigma_i^m, \sigma_{-i})$ under

[21] This is because $v^\theta(\sigma)$ can be written as a weighted average of the action profiles played in equilibrium: $v^\theta(\sigma) = \sum \lambda^\theta_{\sigma,a} g^a$, with $\lambda^\theta_{\sigma,a}$ appropriately taking into account the date at which a is played and the discount factor. The statement means that $\lambda^\theta_{\sigma,a^*}$ must be close to 1.

[22] $\widehat{Q}$ contains $\phi^{-1}_{i,\alpha_i^m}(Q^{\sigma,\varepsilon})$ because by definition of $Q^{\sigma,\varepsilon}$, $Q^{\sigma,\varepsilon} \subset B^m$, and because α_i^m has been chosen so that $\phi^{-1}_{i,\alpha_i^m}(B^m) \subset B^{2m} \subset Q$.

$\widehat{Q}$ also contains $B^m \backslash Q^{\sigma,\eta_0}$ by definition of Q^{σ,η_0} and $\widehat{Q}$.

θ and by σ under θ' coincide except that whenever a^* is played under θ', (σ_i^m, a_{-i}^*) is played under θ. Since $\theta' \in Q^{\sigma,\eta_0}$, that path must prescribe that a^* is played except for a fraction of the time of order η_0. So by construction we have:

$$v_i^\theta(\sigma_i^m, \sigma_{-i}) \geq g_i^\theta(a^*) + \Delta_i^m + O(\eta_0).$$

Since $\sigma \in \mathcal{E}^{\delta,\theta}$, we have:

$$v_i^\theta(\sigma_i^m, \sigma_{-i}) \leq v_i^\theta(\sigma)$$

thus implying:

$$v_i^\theta(\sigma) \geq g_i^\theta(a^*) + \Delta_i^m + O(\eta_0)$$

thus contradicting the premise that a^* is played most of the time under θ. So $\phi_{i,\alpha_i^m}^{-1}(Q^{\sigma,\eta_0}) \cap Q^{\sigma,\eta_0} = \varnothing$ as desired. Q.E.D.