On Failing to Cooperate When Monitoring is Private

Olivier Compte*

C.E.R.A.S. †
Ecole Nationale des Ponts et Chaussées
28 Rue des St-Pères
75007 Paris, France

February 2000

*I am grateful to the referees and the associate editor for their helpful comments. I thank seminar audiences at CORE, Ecole Normale Superieure in Paris and Center (Tilburg) for their comments. I also thank Bernard Caillaud, Philippe Jehiel, Frederic Jouneau, Jean Tirole and Tom Palfrey for helpful discussions and comments.

†URA cnrs 2036; e-mail: compte@enpc.fr

Abstract: I consider a repeated prisoners' dilemma where in each period, each player receives an imperfect private signal about his opponent's current action. I show that when players are patient enough, *any* equilibrium where players use trigger strategies (i.e. do not revert to cooperation once they have started defecting) yields players a value arbitrarily close to the mutual minimax. I also examine the robustness of the result to perturbations of the game.

1 Introduction

Repeated game models have been used in many economic settings to show that repetition may allow economic agents to enforce efficient agreements. Since Friedman (1971), the theory has enlarged the set of conditions that would allow firms to support tacit collusion. Green and Porter (1984) have allowed signals to be imperfectly correlated with actions, and Abreu Pearce and Stacchetti (1990), Fudenberg Levine and Maskin (1994) and Fudenberg and Levine (1994) have provided general tools to compute or characterize the set of equilibrium outcomes. Yet those papers assume that signals are public and they restrict their analysis to public equilibria, that is, equilibria where in equilibrium players condition their actions on public information only. So they cannot deal with the case where the signals that players receive are private. In particular, they cannot address the issues brought up by Stigler (1964). In Stigler, firms do not observe the prices charged by their competitors; they use the demand for their own product as a signal of their competitors' pricing strategy.

This paper does not have the ambition to solve Stigler's problem, but to shed some light on why it may be difficult to support cooperation when only *imperfect private* signals are available. To this end, I consider a repeated prisoners' dilemma where at the end of each period, each player receives an imperfect private signal about his opponent's current action and I restrict attention to equilibria where in equilibrium, players use trigger strategies (i.e. do not revert to cooperation once they have started defecting). Trigger strategies are a natural candidate for supporting cooperation, and although that restriction may diminish cooperation possibilities, it is of interest to examine how those strategies perform depending on whether monitoring is public or private.

When the signals that players receive are public (or perfectly correlated), values strictly above the mutual minimax can be supported if signals are informative enough, and values close to the efficient outcome can be supported if signals are very informative. In contrast, in the polar opposite case where conditional on each action profile played, the signals that players receive are independent, we obtain the following result: *even when signals are very informative, if players are patient enough then all the equilibrium values are arbitrarily close to the mutual minimax.*

Supporting cooperation requires that deviations be detected and then *punished*. When signals are public, punishing a player just requires finding an equilibrium that gives the deviator a lower expected payoff. This task is particularly easy in the prisoner's dilemma since having both players defect for ever is self-enforcing (once it is public that both players are going to behave in this way).

With imperfect private signals, there is no way for players to *coordinate* on such a punishment path: under a full support assumption, the only public event at date t contains the whole set of $t-$histories of signals. So if it is public at date t that both players defect for ever, it must be so after any $t-$history of signals. Hence by backward induction, we would conclude that both players defect from the start of the game.

Yet coordination does not seem necessary to deter players from defecting: When a player, say player 1, defects, his opponent is more likely to receive bad signals; if player 2 receives many bad signals, she may infer that her opponent is likely to be defecting, hence to be defecting *for ever* (since player 1 is supposed to be using a trigger strategy), which should then induce player 2 to switch to defect. So player 1 should expect his defections to be matched eventually, which should give him the incentives to cooperate in the first place.

However it turns out that in equilibrium, trigger strategies do not enable players to sustain cooperation during many periods. Let us briefly explain why.

When monitoring is public, players can coordinate on a simultaneous reversion to defection, and they have a lot of freedom in choosing the circumstances under which they will do so. They may for example decide to check every period the last T signals observed, and switch to defect whenever these T signals are bad. Thus many statistical tests can be implemented, and it can be shown that there exists one that supports cooperation for a reasonable length of time (in expectation).

When monitoring is private, coordinating on a simultaneous reversion to defection is not possible, and incentives to use past observations appropriately have to be provided differently. Two factors constrain the way past observations will be taken into account. *First*, a player will take into account past observations to the extent that they give him information about how his opponent will behave in the future. In a trigger equilibrium, a player (say player 2) will thus use her signals to discriminate between the two events:

(A) : *player* 1 *always cooperated so far*, and
(B) : *player* 1 *started to defect.*

Second, a player will actually be willing to trigger the punishment when the relative weight of the above two events reaches an appropriate threshold, determined by a comparison between the gain from continuing to cooperate vs. defecting under each event. In particular, since the gain from switching to defect will turn out to be small (because long-run defections are detected reasonably fast), players will only switch to defect when they are very pessimistic about their opponent being still cooperating.

These two constraints drastically reduce the number of tests that players would be willing to perform in equilibrium, and we will show that in any candidate equilibrium for which players would not get a payoff close to the mutual minmax, players would tend to wait for a very large number of bad signals before triggering a punishment; and the consequence is that short run deviations would go unpunished.

Players lack the incentives to retaliate promptly to bad signals; so one might hope that by perturbing the game and by assuming that in each period, each player may become with positive probability a type that always defects, players would be more suspicious and react more rapidly to bad signals. Our result however will be shown to be robust to such perturbations.

The remainder of the paper is organized as follows. In Section 2 we describe the model. In Section 3 the main results are presented. After giving some general properties of trigger strategy equilibria (Section 4), we provide the main argument (Section 5). In Section 6, we examine further the difference between public and private monitoring and discuss the robustness of our result. We conclude in Section 7.

2 The Model

The model is a repeated prisoner's dilemma. In each period, player i chooses an action $a_i \in \{C, D\}$. Current expected payoffs are described by the following bi-matrix:

	C	D
C	$1,1$	b,c
D	c,b	$0,0$

where $c > 1$ and $b < 0$

We assume that players discount future payoffs. If $u(t)$ denotes the payoff at t, players compute the average discounted payoff $(1-\delta)\sum_0^\infty \delta^t u(t)$. It will also be convenient to use the discount rate r defined by $\delta = e^{-r}$. We will be interested in the case of very patient players (i.e. $r \to 0$).[1]

We assume that at each date, each player i observes an imperfect private signal $y_i \in Y_i \equiv \{\underline{y}, \overline{y}\}$ about the current action played by his opponent. At each period, the action profile played $a = (a_1, a_2)$ generates a distribution $f(\cdot \mid a) \in \Delta(Y_1 \times Y_2)$ over signal profiles. The marginal over player 1's observation $f_1(\cdot \mid a) \equiv \sum_{y_2} f(\cdot, y_2 \mid a) \in \Delta(Y_1)$ is assumed to

[1] We will often use the following approximation:

$$\text{for } t << 1/r, \qquad 1-\delta^t = 1-e^{-rt} = rt + O((rt)^2)$$

where the term $O(x)$ means that there exists a constant a independent of x such that $\mid O(x) \mid \leq ax$.

have *full support*, and similarly for player 2. One possible interpretation is that payments are stochastic.

For player i, the history of the game at t consists of his own actions and signals and is denoted $h_i^t = \{(a_i^0, y_i^0); \ldots; (a_i^{t-1}, y_i^{t-1})\}$. A *strategy* σ_i is a sequence of maps from private histories to mixed actions. Each *strategy profile* $\sigma = (\sigma_1, \sigma_2)$ induces a distribution over histories in the usual way, as well as a distribution over future payoffs.[2] We denote by $v_i(\sigma)$ the corresponding average discounted payoff. A *Nash equilibrium* is a strategy profile such that each player's strategy is a best response to his opponent's strategy.

In this paper, we will focus on a particular class of Nash equilibria. We say that a Nash equilibrium is a *trigger strategy equilibrium* if each player uses a *trigger strategy*, that is, a strategy such that he never reverts to cooperation once he starts defecting. Formally a *strategy* σ_i *is a trigger strategy* if it satisfies

$$\Pr\{a_i^t = D \mid a_i^{t-1} = D, \sigma_i\} = 1 \tag{1}$$

We denote by $\mathcal{V}(\delta)$ the largest payoff a player may obtain in a *trigger strategy equilibrium*; we say that *cooperation can be supported* when $\lim_{\delta \to 1} \mathcal{V}(\delta) > 0$.

3 Main Results

Our objective is to contrast the public observations case and the private observations case. Assumption 1 and 2 below formalize the two polar cases that we consider. We start with the well-know public monitoring case:

Assumption 1 *$f(\overline{y}, \underline{y} \mid a) = f(\underline{y}, \overline{y} \mid a) = 0$ and we let:*

$$p \equiv f(\overline{y}, \overline{y} \mid CC) \text{ and } q \equiv f(\underline{y}, \underline{y} \mid CD) = f(\underline{y}, \underline{y} \mid DC) < p \tag{2}$$

Under assumption 1, the signals that players receive are perfectly correlated. Since $\overline{y}$ is more likely to occur when the opponent cooperates, we will say that $\overline{y}$ is a *good* signal and that $\underline{y}$ is a *bad* signal. Although restricting to trigger strategy equilibria entails a loss of generality, some level of cooperation can be supported when signals are sufficiently informative. We recall here a standard result by Radner Myerson and Maskin:

Proposition 1 *(Radner Myerson and Maskin 1986) Let $\bar{v} = 1 - \frac{1-p}{p-q}(c-1)$. Under Assumption 1, $\lim_{\delta \to 1} \mathcal{V}(\delta) = \max\{0, \bar{v}\}$.*

[2] Note that because of the full support assumption, all private histories of signals have positive probability.

Note that when p is close to 1, trigger strategies permit players to support almost efficient outcomes. The other case we consider is the following:

Assumption 2 $f(y_1, y_2 \mid a) = f_1(y_1 \mid a) f_2(y_2 \mid a)$, *and we let:*

$$p \equiv f_i(\overline{y} \mid CC) \text{ and } q \equiv f_1(\overline{y} \mid CD) = f_2(\overline{y} \mid DC) < p \quad (3)$$

In other words, conditional on the action profile played, the signals that players receive are independent.

Theorem 1 *Under Assumption 2,* $\lim_{\delta \to 1} \mathcal{V}(\delta) = 0.$

Note that the only informational parameters that we consider are p and q, which describe the informativeness of the signals received by a player who cooperates. We do not need to make any assumption about the informativeness of the signals received by a player who defects. Clearly, if by defecting, a player could perfectly assess the action played by his opponent, it would be hard to prevent that player to play Defect for learning purposes (thus for a short period of time only). Consequently, our result has more power in cases where a player who Defects receives poorly or not informative signals.

As mentioned in the introduction, our result is robust to perturbations of the game. We define a perturbation as follows. We assume that each player i may (exogenously) become a type that always defect. At any date t, player i may either be a rational type, or a type that always defects. We let ζ_i^t denote the probability that player i becomes a type that always defects at date t *given that he was rational at date* $t-1$. A *perturbation* of the game is given by a sequence $\zeta = ((\zeta_1^0, \zeta_2^0), ..., (\zeta_1^t, \zeta_2^t), ...)$ of *exogenous transition probabilities.*[3] We let Λ denote the set of perturbations. A strategy profile σ will now denote the strategies of each player when rational. Under any perturbation $\zeta \in \Lambda$, a strategy profile σ induces as before a distribution over future payoffs. We denote by $v_i(\sigma, \zeta)$ the corresponding average discounted payoff for player i if rational, and denote by $\mathcal{V}(\delta, \zeta)$ the largest payoff a (rational) player may obtain in a trigger strategy equilibrium of the perturbed game. We have:

Theorem 2 *Under Assumption 2,* $\lim_{\delta \to 1} \sup_{\zeta \in \Lambda} \mathcal{V}(\delta, \zeta) = 0.$

4 Preliminaries

Our purpose in this section is to derive properties shared by all trigger strategy equilibria. In the main text, we will focus on the case where the game is not perturbed. We leave the treatment of the perturbed game to Appendix B. We start with some notation.

[3]Exogenous transition probabilities are assumed to be independent accross players.

4.1 Notation

Transition probabilities: A realization is an infinite path of actions and signals. To each realization and for each player j, we may associate the date $\widetilde{t}_j \in \mathcal{N} \cup \{+\infty\}$ at which player j switches to defect for ever.[4] Each strategy profile σ induces a probability distribution over realizations, hence a distribution over dates $\widetilde{t}_j$.

A trigger strategy σ_j for player j uniquely defines a sequence of *transition probabilities* $(\varepsilon_j^0, \ldots, \varepsilon_j^t, \ldots)$ where ε_j^t denotes the (conditional) probability that player j switches to defect at t given that both players have always cooperated so far. Formally, let σ^C denote the strategy that consists in cooperating at all dates:

$$\varepsilon_j^t \equiv \mathrm{Pr}_{\sigma_i=\sigma^C,\sigma_j}\{\widetilde{t}_j = t \mid \widetilde{t}_j \geq t\}.$$

Note that the transition probability ε_j^t is computed ex ante. Yet because of the independence assumption, conditioning on the signals received before t by player i would not affect the computation. Consequently, if player i has cooperated until date t, then the probability ε_j^t is also player i's belief *at date* t about player j's behavior.

Likelihood ratios: Consider a *trigger strategy* σ_j played by player j. For any $t-$history of actions and signals h_i^t for player i, we denote by $\mathbf{a}[h_i^t]$ the path of actions followed by player i up to date t, and by $p^t(h_i^t)$ the probability that player j cooperates at t:

$$p^t(h_i^t) \equiv \mathrm{Pr}_{\mathbf{a}[h_i^t],\sigma_j}\{a_j^t = C \mid h_i^t\},$$

and we define player i's *likelihood ratio* that player j defects at t by

$$\mathbf{l}_i^t[h_i^t] \equiv \frac{1-p^t(h_i^t)}{p^t(h_i^t)}.$$

The higher the likelihood ratio, the more pessimistic player i is about player j being still cooperating.

Continuation payoffs and value decomposition: At date t, after the history h_i^t, we may compute the *continuation payoff* $v_i^t(\eta, \sigma_j \mid h_i^t)$ obtained by player i when he follows the strategy η from date t on. Since we focus on trigger strategy equilibria, we will be particularly interested in the private histories for which player i always cooperated. To avoid confusion, we denote by $\widetilde{h}_i^t = \{(C, y_i^0); \ldots; (C, y_i^{t-1})\}$ a private history for player i during which party i always cooperated and by $\widetilde{H}_i^t$ the set of such histories.

[4] The date $\widetilde{t}_j$ is thus the first date t for which $a_j^s = D$ for all $s \geq t$. And when no such date exists, $\widetilde{t}_j = +\infty$.

In what follows, we choose a private history $\widetilde{h}_i^t \in \widetilde{H}_i^t$ and decompose the continuation payoff $v_i^t(\eta, \sigma_j \mid \widetilde{h}_i^t)$ into two terms, depending on whether player j is still cooperating or not. Under the event where player j cooperates at t (this event has probability $p^t(\widetilde{h}_i^t)$), both players have been cooperating so far and because of the conditional independence assumption, the distribution over player j's private history does not depend on player i's signals; hence player i's expected payoff depends only on η and σ_j, and we denote it by $\bar{v}_i^t(\eta, \sigma_j)$. For any private history $\widetilde{h}_i^t \in \widetilde{H}_i^t$, we have:

$$\bar{v}_i^t(\eta, \sigma_j) \equiv v_i^t(\eta, \sigma_j \mid \widetilde{h}_i^t, a_j^t = C),$$

Under the event where player j defects at t (this event has probability $1 - p^t(\widetilde{h}_i^t)$), player j defects for ever (because σ_j is assumed to be a trigger strategy); we denote by σ^D that strategy. For any private history $\widetilde{h}_i^t \in \widetilde{H}_i^t$, we may thus *decompose* continuation payoffs as follows:

$$v_i^t(\eta, \sigma_j \mid \widetilde{h}_i^t) = p^t(\widetilde{h}_i^t)\bar{v}_i^t(\eta, \sigma_j) + (1 - p^t(\widetilde{h}_i^t))v_i^t(\eta, \sigma^D) \tag{4}$$

Continuation values and the cautious strategies

For any trigger strategy σ_j, we let $\bar{v}_i^t(\sigma_j)$ denote the *continuation value* obtained by player i from date t on when i) player j follows σ_j, ii) player j still cooperates at t, iii) player i switches to defect *at the same date* as player j does. That is, player i obtains a payoff equal to 1 from date t until date $\widetilde{t}_j - 1$, and a payoff equal to 0 afterwards:

$$\bar{v}_i^t(\sigma_j) \equiv E_{\sigma_i = \sigma^C, \sigma_j}[1 - e^{-r(\widetilde{t}_j - t)} \mid \widetilde{t}_j > t]. \tag{5}$$

Because signals are imperfect, player i cannot switch to defect exactly at the same date as player j does. In particular, player i may receive bad signals even when player j cooperates, in which case player i would switch to defect *before* player j does; or player i may get good signals even when player j defects, in which case player i would switch to defect *after* player j does. Yet if player i tests player j's behavior during a large (but not too large) number of periods before deciding whether to switch to defect, both types of mistakes will be unlikely to occur for very long, and player i should be able to secure a payoff close to $\bar{v}_i^t(\sigma_j)$ in the event player j still cooperates at t.

Formally, for any scalar $\beta > 0$, we define the *cautious strategy* η_β^* for player i as follows. For each signal $y_i \in \{\bar{y}, \underline{y}\}$, we define the likelihood ratio

$$\alpha(y_i) \equiv \frac{\Pr_{a_i = C, a_j = D}(y_i)}{\Pr_{C,C}(y_i)} = \frac{f(y_i \mid CD)}{f(y_i \mid CC)}.$$

For any date t, consider the last $\beta \mid \ln r \mid$ dates where party i cooperated, and let $\mathcal{T}_\beta^t$ denote the set of theses dates. Party i cooperates at t if $t \leq \beta \mid \ln r \mid$ or if

$$\prod_{s \in \mathcal{T}_\beta^t} \alpha(y_i^s) \leq 1,$$

and he defects for ever otherwise. We will check in Lemma 5 that for an appropriately chosen scalar β, party i may secure payoff close to $\bar{v}_i^t(\sigma_j)$ (in the event player j still cooperates at t).

4.2 Properties of trigger equilibria

In what follows, we denote by $\mathcal{E}(r)$ the *set of trigger equilibria* when the discount rate is equal to r. We establish properties of these trigger equilibria. Note that all the proofs of the Lemma in this section are relegated to the Appendices. Intuition for each result is provided in the main text.

Existence of thresholds: We first establish that in a trigger equilibrium, a player switches to defect when the probability that his opponent defects crosses a given threshold. We have:

Lemma 1 *Let $\overline{L} = \frac{c}{-b(1-\delta)}$. In any trigger strategy equilibrium $\sigma \in \mathcal{E}(r)$, and for each player i, there exists a sequence of thresholds $(\bar{l}_i^0, \ldots \bar{l}_i^t, \ldots) \in [0, \overline{L}]^\infty$ such that player i continues to cooperate if (and only if) $\mathbf{l}_i^t[\widetilde{h}_i^t] < (\leq) \bar{l}_i^t$. The threshold $\bar{l}_i^t$ satisfies:*

$$\bar{l}_i^t \equiv \max\{ \sup_{\eta_i \neq \sigma^D} \frac{\bar{v}_i^t(\eta_i, \sigma_j) - \bar{v}_i^t(\sigma^D, \sigma_j)}{-v_i(\eta_i, \sigma^D)}, 0\}$$

Proof. See Appendix B. ■

Lemma 1 is a direct consequence of equation (4): for player i, only two events matter for future payoffs: either (A): *player j still cooperates*, or (B): *player j starts or has started to defect.* Whether player i should continue to cooperate or switch to defect depends on the relative weight of those two events.

Dynamic of beliefs: To pursue the analysis of equilibrium strategies, we describe how the likelihood ratios $\mathbf{l}_i^t[\widetilde{h}_i^t]$ evolve over time depending on the signals received. The next Lemma is a straightforward application of Bayes law:

Lemma 2 *Consider $\sigma \in \mathcal{E}(r)$ and $\widetilde{h}_i^t \in \widetilde{H}_i^t$. Let $l_i^t = \mathbf{l}_i^t[\widetilde{h}_i^t]$. If player i receives the signal y_i^t at date t, we have:*

$$\mathbf{l}_i^{t+1}[\widetilde{h}_i^{t+1}] = \frac{\alpha(y_i^t)}{1-\varepsilon_j^{t+1}} l_i^t + \frac{\varepsilon_j^{t+1}}{1-\varepsilon_j^{t+1}} \tag{6}$$

Thus when the signal $\underline{y}$ is received, this is bad news and the likelihood ratio increases (because $\alpha(\underline{y}) = \frac{1-q}{1-p} > 1$). And when the signal $\overline{y}$ is received, this is good news and the likelihood ratio decreases (unless ε_j^{t+1} is too large). As a benchmark case, assume that player 1 defects with positive probability ε^0 at date 0, and that $\varepsilon_1^t = 0$ afterwards. At $t = 0$, player 2's likelihood ratio is $l_2^0 = \frac{\varepsilon^0}{1-\varepsilon^0}$. At date t, after the sequence of signals $(y_2^0, ... y_2^{t-1})$ has been received, we have:

$$l_2^t = l_2^0 \prod_{0 \leq s < t} \alpha(y_2^s) \tag{7}$$

This is a standard learning problem and player 2 tends to learn what player 1 is doing: if player 1 cooperates, l_2^t almost surely tends to 0, and if player 1 defects, l_2^t almost surely tends to ∞.

When player 1 switches to Defect with positive probability after the first date, then player 2 has to take in account the fact that player 1 may have switched to Defect recently, which explains (6). Nevertheless, under the event where likelihood ratios are large compared to transition probabilities, the dynamic of beliefs is close to that given by (7). It is thus useful at this stage to derive more explicitly the rate at which the product $\prod_{t=0}^{T} \alpha(y_2^t)$ tends to 0 or to ∞, depending on the strategy of player 1.

Let $m = E_{C,C}[\ln \alpha(y_2)]$ (< 0) and $m^D = E_{D,C}[\ln \alpha(y_2)]$ (> 0). For any $\tau \geq 1$, we let σ^τ *denote the strategy where party* 1 *defects once every* τ *period.* Choose τ^* such that

$$\frac{m^D}{\tau^*} + (1 - \frac{1}{\tau^*})m = \frac{m}{2}.$$

The following result is standard:

Lemma 3 *There exists β such that $\forall r, T \geq \beta \mid \ln r \mid, \tau \geq \tau^*$*

$$\Pr_{\sigma^\tau, \sigma^C}\{\prod_{t=1}^{T} \alpha(y_2^t) \geq e^{-\frac{|m|T}{4}}\} \leq r^4, \text{ and} \tag{8}$$

$$\Pr_{\sigma^D, \sigma^C}\{\prod_{t=1}^{T} \alpha(y_2^t) \leq e^{\frac{m^D T}{2}}\} \leq r^4. \tag{9}$$

Proof. See Appendix A ■

The interpretation of (8) is that even if party 1 defects a substantial fraction of the time, player 2's likelihood ratio will nevertheless tend to 0 at an exponential rate. The interpretation of (9) is that when player 1 defects, player 2's likelihood ratio tends to $+\infty$ at an exponential rate.

Lower bound on thresholds: As explained before, we may associate to any given trigger equilibrium $\sigma \in \mathcal{E}(r)$ a sequence of transition probabilities $\{(\varepsilon_1^t, \varepsilon_2^t)\}_{t\geq 0}$, a sequence of continuation values $\{(\bar{v}_1^t, \bar{v}_2^t)\}_{t\geq 0}$ (where $\bar{v}_i^t \equiv \bar{v}_i^t(\sigma_j)$) and a sequence of thresholds $\{(\bar{l}_1^t, \bar{l}_2^t)\}_{t\geq 0}$. The following result establishes a key relationship between these various sequences.

Proposition 2 *Choose β as in Lemma 3. There exists r_0, a, d such that $\forall r \leq r_0$, $\forall \sigma \in \mathcal{E}(r)$, $\forall t$,*

$$\bar{l}_i^t \geq \frac{a\bar{v}_i^t - dr \mid \ln r\varepsilon^t \mid}{r \mid \ln r \mid} \qquad \forall i$$

where

$$\varepsilon^t = \max_{s, t-\beta|\ln r| \leq s \leq t} \max\{\varepsilon_1^s, \varepsilon_2^s\}$$

Note that the strength of Proposition 2 lies in the fact that the constants a and d can be chosen independently of r and of the equilibrium considered. Proposition 2 says that except at dates where the transition probabilities are small or at dates where $\bar{v}_i^t$ is small (which means that player j will soon switch to defect), the threshold $\bar{l}_i^t$ must be high (that is, comparable to $\frac{1}{r|\ln r|}$).

Proof. The proof exploits the definition of $\bar{l}_i^t$ given in Lemma 1 and computes a lower bound on $\bar{l}_i^t$ by looking at the (relative) performance of the cautious strategy η_β^* (where β is defined as in Lemma 3) and of the strategy σ^D. More precisely, the proof relies on the following observations:

Observation 1. The payoff $\bar{v}_i^t(\sigma^D, \sigma_j)$ is at most comparable to $r \mid \ln(\varepsilon_i^t r) \mid$. Indeed, when player i starts to defect at t, the likelihood ratio of his opponent grows exponentially (because of Lemma 3). Since $l_j^t \geq \varepsilon_i^t$ (by Lemma 2), player j should reach the maximum threshold $\overline{L}$ (hence switch to defect) in a number of periods comparable to $\mid \ln(\varepsilon_i^t r) \mid$. In Appendix B, we actually prove a stronger result:

Lemma 4 *Choose β as in Lemma 3. There exist $d, r_0 > 0$ such that $\forall r \leq r_0$, $\forall \sigma \in \mathcal{E}(r)$, $\forall t$:*

$$\bar{v}_i^t(\sigma^D, \sigma_j) \leq dr \mid \ln r\varepsilon^t \mid$$

where $\varepsilon^t = \max_{s, t-\beta|\ln r| \leq s \leq t} \max\{\varepsilon_1^s, \varepsilon_2^s\}$

Observation 2. As mentioned at the end of Section 4.1., party i should be able to secure a payoff close to $\bar{v}_i^t(\sigma_j)$ by following a cautious strategy η^*_β. Formally, we have:

Lemma 5 *Choose β as in Lemma 3 and let $\bar{T} = \frac{|\ln r|}{r}$. There exist $a, r_0, \forall r \leq r_0$, $\forall \sigma_j$, $\bar{v}_i^t(\eta^*_\beta, \sigma_j) \geq \bar{v}_i^t(\sigma_j) - dr \mid \ln r \mid$ and $v_i^t(\eta^*_\beta, \sigma^D) \geq -dr \mid \ln r \mid$*

Lemma 5 is a straightforward corollary of Lemma 3 (see Appendix B). The intuition is that when player i follows the cautious strategy η^*_β, he is very unlikely to switch to defect before player j does (thanks to inequality (8)). And he is also very unlikely to switch to defect more than $2\beta \mid \ln r \mid$ periods after player j does (thanks to inequality (9).

Combining Lemma 2, 4 and 5 yields the desired lower bound. ■

Value of the game and continuation values. To any trigger equilibrium $\sigma = (\sigma_1, \sigma_2)$, we may associate the continuation values $(\bar{v}_1^0, \bar{v}_2^0) \equiv (\bar{v}_1^0(\sigma_2), \bar{v}_2^0(\sigma_1))$ and the initial transition probabilities ε_i^0. To conclude this Section, we relate the value of the game $v_i(\sigma)$ to $\bar{v}_1^0$ and ε_2^0.

Lemma 6 $\exists a, r_0 > 0, \forall r \leq r_0$, $\forall \sigma \in \mathcal{E}(r), \forall i$

$$\mid v_i(\sigma) - (1 - \varepsilon_2^0)\bar{v}_1^0 \mid \quad \leq \quad ar \mid \ln r \mid \tag{10}$$

Note that inequality (10) implies that $v_1(\sigma)$ cannot differ from $v_2(\sigma)$ by more than $ar \mid \ln r \mid$. The intuition is that in equilibrium, a player cannot expect to gain much by deviating before the other does: either this event has a small probability, or his opponent will rapidly switch to defect, because his likelihood ratio will rapidly reach the threshold $\bar{L}$.

5 The Main Argument

In this Section, we fix $v > 0$ and consider *the set of trigger equilibria $\mathcal{E}(v, r)$ yielding one player a payoff at least equal to v.* Assuming that this set is not empty, we fix a trigger equilibrium $\sigma \in \mathcal{E}(v, r)$ and compute a lower bound on the payoff obtained by party 1 when he deviates. The class of deviations we consider is the following:

Definition 1 *(The deviation $\sigma^{\tau,\beta}$ and the strategy σ^τ). Let T^* be such that $1 - e^{-rT^*} = \frac{v}{2}$. Let K be such that $K\tau \leq \frac{T^*}{2} \leq (K+1)\tau$, and choose β as in Lemma 3. The* strategy *$\sigma^{\tau,\beta}$ is defined by:*

i) *defect at all dates $t = k\tau$, with $1 \leq k \leq K$,*

ii) *at any other date, follow the cautious strategy* η^*_β.

To define the strategy σ^τ*, we replace condition* **ii)** *by:*

ii′) *cooperate at any other date.*

The strategy $\sigma^{\tau,\beta}$ thus consists in defecting once every τ periods until date $T^*/2$, and following the cautious strategy η^*_β at other dates. Note that when he follows $\sigma^{\tau,\beta}$, party 1 is very unlikely to switch to defect for ever before player 2 does.[5] And under the event where player 1 switches to defect after player 2 does, the strategy actually followed by player 1 (until player 2 switches) is σ^τ. For this reason, understanding the effect of the *strategy* σ^τ on player 2's behavior will be key.

5.1 The effect of σ^τ on party 2's behavior.

As explained earlier, the trigger equilibrium considered σ defines a sequence of transition probabilities $\{(\varepsilon^t_1, \varepsilon^t_2)\}_{t\geq 0}$, a sequence of continuation values $\{(\bar{v}^t_1, \bar{v}^t_2)\}_{t\geq 0}$ and a sequence of thresholds $\{(\bar{l}^t_1, \bar{l}^t_2)\}_{t\geq 0}$. It also defines, thanks to Lemma 2, each player's current likelihood l^t_i as a function of the history of signals received (on paths where player i has not switched yet to defect).

Given these sequences and the likelihood ratio function $\mathbf{l}^t_i$, we wish to assess how the probability that party 2 switches to Defect is affected when party 1 follows σ^τ instead of σ_1. To this end, we define the probability $\pi^t(\tau)$ that party 2's likelihood ratio reaches $\bar{l}^t_2$ for the first time at t when party 1 follows σ^τ (instead of σ_1) and when player 2 cooperates:[6]

$$\pi^t(\tau) \equiv \Pr_{\sigma^\tau,\sigma^C}\{\mathbf{l}^t_2[h^t_2,\sigma_1] \geq \bar{l}^t_2, \mathbf{l}^s_2[h^s_2,\sigma_1] < \bar{l}^s_2 \ \forall s < t\} \tag{11}$$

Our objective (in the next section) is to bound these probabilities: if these probabilities are small, then player 1 will not be punished harshly from deviating, and the deviation will be profitable. The following Lemma gives a lower bound on the gain from following $\sigma^{\tau,\beta}$ when $\pi^t(\tau)$ is small during many periods.

Lemma 7 *Let* $G(\tau) = \frac{v^2}{10\tau}(c-1)$*. Choose* $\sigma \in \mathcal{E}(v,r)$*. If there exists* $T \geq T^*/2$ *and* $\bar{\tau}$ *such that* $\sum_{t<T} \pi^t(\tau) \leq r^{1/6}$ *for all* $\tau \leq \bar{\tau}$*, then we must have:*

$$v_1(\sigma^{\tau,\beta},\sigma_2) - v_1(\sigma) \geq G(\bar{\tau}) - e^{-rT}\bar{v}^T_1(\sigma_2)$$

[5] This follows from Lemma 3. Player 1 switches to defect before player 1 does with probability at most equal to $r^4\bar{T}$ if $\widetilde{t}_1 \leq \bar{T} = \frac{|\ln r|}{r}$. And what happens after date $\bar{T}$ has a negligible impact on player 1's discounted payoff when r is small, as $e^{-r\bar{T}} = r$.

[6] We use the notation $\mathbf{l}^t_2[h^t_2,\sigma_1]$ in the definition of $\pi^t(\tau)$ to emphasize the fact that player 2 computes the likelihood ratio l^t_2 assuming that player 1 follows his supposed equilibrium strategy σ_1; whereas the (true) distribution over histories received by player 2 is determined by the strategy profile (σ^τ, σ^C).

In particular, if T is very large, e^{-rT} is small and since $\bar{v}_1^T(\sigma_2)$ cannot exceed 1, the deviation is profitable. The term $G(\tau)$ corresponds to a lower bound on the extra gain that player 1 obtains during the first $T^*/2$ periods by defecting once every τ periods instead of cooperating. The term $e^{-rT}\bar{v}_1^T(\sigma_2)$ corresponds to an upperbound on the discounted loss that player 1 may incur after date T, due to the reaction of player 2 to the deviation of player 1. The intuition for the result is that when players follow $(\sigma^{\tau,\beta}, \sigma_2)$, they actually follow the strategy profile (σ^τ, σ^C) until date T with probability close to 1. Details are in Appendix C.

We will prove the following Proposition:

Proposition 3 *There exists r_0 such that $\forall r \leq r_0$, $\forall \sigma \in \mathcal{E}(v, r)$, we may find $T^{**} \geq T^*/2$ and τ^{**} such that*

$$\sum_{t<T^{**}} \pi^t(\tau) \leq r^{1/6} \text{ for all } \tau \leq \tau^{**}, \text{ and} \tag{12}$$

$$e^{-rT^{**}}\bar{v}_1^{T^{**}}(\sigma_2) \leq \frac{G(\tau^{**})}{2} \tag{13}$$

Proposition 3 thus says that for any trigger equilibrium $\sigma \in \mathcal{E}(v, r)$, we may find a strategy $\sigma^{\tau^{**}}$ and a date T^{**} such that i) before T^{**}, the effect of $\sigma^{\tau^{**}}$ on player 2's behavior is negligible, and ii) after T^{**}, the discounted expected payoff that player 1 would have gained by following σ_1 from the start is smaller than the gain obtained by following $\sigma^{\tau^{**}}$ up to date T^{**}.

Proposition 3 and Lemma 7 thus imply that $\sigma^{\tau^{**},\beta}$ is a profitable deviation, which concludes the proof of our main result. The rest of the Section will be devoted to the proof of Proposition 3.

Comment 1: One corollary of Proposition 3 is that the rate at which player 2 (voluntarily) switches to defect when player 1 cooperates must be small: over a length of time comparable to $1/r$ (because $T^*/2$ is comparable to $1/r$), the total probability that player 2 switches to defect is at most equal to $r^{1/6}$. Hence the average transition rate is much smaller than r. One interpretation is that player 2 waits for many bad signals before switching to defect. And this is precisely why short run deviations turn out be profitable.

Comment 2: The probability that the likelihood ratio reaches a threshold $\bar{l}_i^t$ should become smaller when the threshold becomes larger. In the next Section, we derive a key statistical result that will make precise the relationship between the thresholds and the probabilities to switch to defect. Proposition 2, which has derived a lower bound on these thresholds $\bar{l}_i^t$, will then permit us to derive an upperbound on the probabilities $\pi^t(\tau)$.

5.2 A key statistical result.

Consider any given trigger strategy σ_1 and the sequence $\{\varepsilon_1^t\}_{t\geq 1}$ associated to it, and assume that player 2's likelihood ratio evolves according to (6) (that is, player 2 believes that player 1 is following σ_1). We wish to assess how player 2 behaves when player 1 actually follows σ^τ instead of σ_1. To this end, we let[7]

$$\pi^{t,T}(\tau, \bar{l}, l^0) \equiv \max_{0\leq s\leq T} \Pr_{\sigma^\tau,\sigma^C}\{\mathbf{l}_2^t[h_2^t, \sigma_1] \geq \bar{l} \mid \mathbf{l}_2^{t-s}[h_2^{t-s}, \sigma_1] = l^0\}$$

denote the maximum probability that l_2^t reaches $\bar{l}$ given that player 2's likelihood ratio was equal to l^0 at some earlier date $t-s$, $s \leq T$ (given that players follow the strategy profile (σ^τ, σ^C)). We have:

Lemma 8 *Define τ^* as in Lemma 3 and let $\rho^{t,T} = \prod_{0\leq s<T}(1-\varepsilon_1^{t-s})$.[8] There exist $\lambda > 0$ and $r_0 > 0$ such that: $\forall \tau \geq 2\tau^*$, $\forall r < r_0$,*

$$\varepsilon_1^{t-s} \leq \varepsilon \ \forall s \leq T \Rightarrow \pi^{t,T}(\tau, \bar{l}, l^0) \leq \left[\frac{\max(\varepsilon \mid \ln r \mid, l^0)}{\bar{l}\rho^{t,T}}\right]^{1-\lambda[\frac{1}{\tau}+\frac{1}{|\ln r|}]}$$

This is a slight extension of a familiar result, which states that when $\varepsilon = 0$, $\pi^{t,T}(\infty, l, l^0) \leq \frac{l^0}{l}$. Indeed, if $\tau = \infty$, then $\sigma^\tau = \sigma^C$, and if $\varepsilon = 0$, the likelihood ratio l_2^t is a martingale, that is, $E_{\sigma^C,\sigma^C}[l_2^t \mid l_2^0 = l^0] = l^0$. Since l_2^t is positive, $E_{\sigma^C,\sigma^C}[l_2^t \mid l_2^0 = l^0] \geq l \Pr_{\sigma^C,\sigma^C}[l_2^t \geq l \mid l_2^0 = l^0]$, and we obtain the desired upperbound. When ε is different from 0, or when player 1 defects from time to time, the likelihood ratio l_2^t is not a martingale anymore, but the drift is small when l_2^t becomes large compared to ε (for example when $\frac{\varepsilon}{l_2^t} \leq \frac{1}{|\ln r|}$) and when the frequency of defections ($\frac{1}{\tau}$) is small. And as $\frac{1}{|\ln r|}$ and $\frac{1}{\tau}$ get small, the bound on $\pi^{t,T}(\tau, \bar{l}, l^0)$ gets comparable to the case where l_2^t follows a martingale. The complete proof is in Appendix A.

5.3 A simple case.

Before proceeding to the general case, we wish to consider a simpler case. In this subsection, we restrict our attention to (hypothetical) symmetric equilibria in $\mathcal{E}(v, r)$ for which

i) the transition probabilities ε_i^t satisfy $\varepsilon_i^t = \varepsilon \ \forall i, \forall t$, for some $\varepsilon > 0$, and

ii) the continuation values $\bar{v}_i^t$ satisfy $\bar{v}_i^t \geq \gamma v \ \forall i, \forall t$ for some $\gamma > 0$.

[7] As for the definition of $\pi^t(\tau)$, we write $\mathbf{l}_2^t[h_2^t, \sigma_1]$ to indicate that the likehood ratio is computed assuming that player 1 follows σ_1. The (true) distribution over histories is the one induced by (σ^τ, σ^C).

[8] Note that $\rho^{t,T} = \Pr_{\sigma_1,\sigma^C}\{\widetilde{t}_1 > t \mid \widetilde{t}_1 \geq t - T + 1\}$. It is the probability that player 1 switches to defect after t given that he did not switch to defect before $t - T + 1$, when players follow (σ_1, σ^C).

We will consider the relevant case where ε is not too large, say $\varepsilon \leq r^{1/2}$.[9]

We will prove that Proposition 3 holds with $T^{**} = \bar{T}$, where $e^{-r\bar{T}} = r$.

Proof of Proposition 3 (simple case): We choose $\underline{\varepsilon}$ and $\bar{l}$ such that

$$| \ln r\underline{\varepsilon} |= \frac{a\gamma v}{2dr}, \text{ and } \bar{l} = \frac{a\gamma v}{2r \mid \ln r \mid}$$

We distinguish these two cases:

Case 1: $\varepsilon \geq \underline{\varepsilon}$.

Then Proposition 2 implies that $\bar{l}_2^t \geq \bar{l}$. For any sequence of signals driving l_2^t above $\bar{l}_2^t$, there exists a last date t_0 such that $l_2^{t_0}(h_2^{t_0}) < \varepsilon \mid \ln r \mid$. Choose $T_0 = \beta \mid \ln r \mid$ and consider the events $E = \{t_0 \geq t - T_0\}$ and $F = \{t_0 < t - T_0\}$. Under event E, we have

$$\Pr_{\sigma^\tau,\sigma^C}\{l_2^t \geq \bar{l} \mid E\} \leq \pi^{t,T_0}(\tau, \bar{l}, \varepsilon \mid \ln r \mid) \tag{14}$$

by definition of π^{t,T_0}. Since $\varepsilon \leq r^{1/2}$, Lemma 8 implies that the right hand side of (14) is comparable to $[r^{3/2} \mid \ln r \mid^2]^{1-\lambda/\tau}$, which is smaller than $r^{5/4}$ for all $r \leq r_0$ if τ large enough and r_0 small enough.

Consider now event F. First observe that Equation (6), which defines the likelihood ratio as a function of the signals received (see Lemma 2), implies:

$$l_2^{t-s+1} \leq l_2^{t-s}\alpha(y_2^{t-s+1})\frac{1 + \frac{1}{\alpha(\bar{y})|\ln r|}}{1 - \varepsilon_1^{t-s+1}}$$

hence,

$$l_2^t \leq l_2^{t-T_0} \prod_{0 \leq s < T_0} \alpha(y_2^{t-s})\frac{(1 + \frac{1}{\alpha(\bar{y})|\ln r|})^{T_0}}{\rho^{t,T_0}}.$$

Since T_0 is comparable to $\mid \ln r \mid$, and since $\varepsilon_1 \leq r^{1/2}$, the ratio on the right hand side is bounded by a constant independent of r. The product $\prod_{0 \leq s < T_0} \alpha(y_2^{t-s})$ gets below $r^{-\frac{|m|\beta}{4}}$ with probability $1 - r^4$ by Lemma 3. If $l_2^{t-T_0} \leq \bar{L}$, then $l_2^t \geq \bar{l}$ requires $\frac{l_2^t}{l_2^{t-T_0}} \geq \frac{\bar{l}}{\bar{L}}$. This last term is comparable to $\frac{1}{|\ln r|}$, hence it is negligible compared to $r^{-\frac{|m|\beta}{4}}$ (for r small enough). Thus we obtain

$$\Pr_{\sigma^\tau,\sigma^C}\{l_2^t \geq \bar{l} \mid F, l_2^{t-T_0} \leq \bar{L}\} \leq r^4 \tag{15}$$

Combining (14) and (15) implies $\sum_{t \leq \bar{T}} \pi^t(\tau) \leq r^4\bar{T}$, which is smaller than $r^{1/6}$ for all $r \leq r_0$ if r_0 is small enough.

Case 2: $\varepsilon < \underline{\varepsilon}$.

[9] If $\varepsilon > r^{1/2}$, cooperation collapses with probability close to 1 in a number of periods small compared to $1/r$, and equilibrium payoffs are then close to 0.

Let $z = (y_2^{k\tau})_{1\leq k\leq K}$ denote a sequence of signals received by player 2 at the dates where player 1 defects. Let Z be the set of such sequences. Since σ^τ only differ from σ^C at these dates, we have:

$$\begin{aligned}\sum_{s\leq t}\pi^s(\tau) &= \Pr_{\sigma^\tau,\sigma^C}\{\exists s\leq t, l_2^s \geq \bar{l}_2^s\} \\ &\leq \sum_{z\in Z}\Pr_{\sigma^C,\sigma^C}\{\exists s\leq t, l_2^s\geq \bar{l}_2^s \mid z\}\Pr_{\sigma^\tau,\sigma^C}\{z\} \\ &\leq \max_{z\in Z}\Pr_{\sigma^C,\sigma^C}\{\exists s\leq t, l_2^s\geq \bar{l}_2^s \mid z\}\end{aligned}$$

Now let $\rho = \min(p, 1-p)$. Since $\Pr_{\sigma^C,\sigma^C}(z) \geq \rho^K$ and since $\Pr_{\sigma^C,\sigma^C}\{\exists s \leq t, l_2^s \geq \bar{l}_2^s\} \leq t\underline{\varepsilon}$, we obtain:

$$\sum_{s\leq \bar{T}}\pi^s(\tau) \leq \frac{\bar{T}\underline{\varepsilon}}{\rho^K}.$$

Finally, observe that $K \leq \frac{T^*}{2\tau}+1$, $T^* = \frac{|\ln(1-v/2)|}{r}$, and $|\ln r\underline{\varepsilon}| = \frac{a\gamma v}{2dr}$. So for τ large enough (yet chosen independently of r), the ratio $\frac{\bar{T}\underline{\varepsilon}}{\rho^K}$ is smaller than $\frac{|\ln r|}{r^2}e^{-\frac{a\gamma v}{4dr}}$, hence smaller than $r^{1/6}$ for all $r \leq r_0$ if r_0 is small enough. ■

5.4 The general case.

Outline of the proof Our proof has two main parts. Part 1 focuses on the short run effect of the deviation. Part 2 focuses on the long-run effect of the deviation (short-run and long-run will be defined shortly).

In the simple case analyzed in the previous section, either transition probabilities were very small (smaller than $\bar{\varepsilon}$) or the thresholds were large (larger than $\bar{l}$). A key reason as to why this was the case is that continuation values $\bar{v}_2^t$ did not vanish over time (they remained at least equal to γv for some $\gamma > 0$). The main insight of part 1 is that in the short run, continuation values do not vanish (otherwise we would not have $v_i(\sigma) = v$), and as a result (so we will show) the analysis of the simple case carries over to the general case.

In the long run, the analysis of the simple case may not be applicable. However, when player 1 follows σ^τ, he stops defecting after $T^*/2$. Therefore the signals that player 2 receives after that date should tend to convince her that party 1 has not started to defect yet. And after some time, recent signals should bear much more weight on player 2's behavior than the signals received before $T^*/2$. Part 2 will formalize this simple intuition.

Part 1: Short run effect. For any $\gamma \in (0, 1/2]$, we define a critical date:[10]

$$T(\gamma) = \max\{t \mid t \leq \bar{T}, \text{ and } \forall s \leq t, \forall i, \bar{v}_i^s \geq \gamma v\},$$

which means that before date $T(\gamma)$, continuation values remain larger than γv. Observe that (for r small enough), $T(\frac{1}{2}) \geq T^*$.[11]

For any $\nu > 0$, we let $\varepsilon(\nu) = \exp -\frac{\nu}{r}$ and define the critical date

$$T_1(\nu) = \min\{t \geq \frac{3T^*}{4} \mid \exists i, \varepsilon_i^t \geq \varepsilon(\nu)\},$$

which means that between date $\frac{3T^*}{4}$ and date $T_1(\nu)$, transition probabilities are (extremely) small for both players. Finally, we let

$$\bar{T}(\gamma, \nu) = \max\{T_1(\nu), T(\gamma)\}$$

That is, before date $\bar{T}(\gamma, \nu)$, either continuation values remain above γv or transition probabilities are smaller than $\varepsilon(\nu)$. The next Proposition directly builds on the analysis of the simple case:

Proposition 4 *Let $\gamma \in (0, 1/2]$ and $\nu > 0$. There exists $r_{\gamma,\nu}, \tau_{\gamma,\nu} > 0$ such that $\forall r \leq r_{\gamma,\nu}$, $\forall \sigma \in \mathcal{E}(v, r)$, $\forall \tau \geq \tau_{\gamma,\nu}$,*

$$\sum_{t \leq \bar{T}(\gamma,\nu)} \pi^t(\tau) \leq r^{1/5}$$

Proof. Compared to Proposition 3 where transition probabilities were bounded above by $r^{1/2}$, there may now be dates where transition probabilities are large. However, this new difficulty will be easily dealt with because there cannot be many such dates. We choose β as in Lemma 3, let $T_0 = \beta \mid \ln r \mid$, and define the set of exceptional dates $\mathcal{S}$ by:

$$\mathcal{S} = \{t \leq \bar{T} \mid \exists s, i \text{ such that } t - T_0 \leq s \leq t \text{ and } \varepsilon_i^s > r^{1/2}\}. \tag{16}$$

Finally, we choose $\underline{\varepsilon}$ and $\bar{l}$ as in the simple case: $\mid \ln r \underline{\varepsilon} \mid = \frac{a\gamma v}{2dr}$ and $\bar{l} = \frac{a\gamma v}{2r|\ln r|}$. Recall that $\varepsilon^t = \max_{i,s,t-T_0 \leq s \leq t} \varepsilon_i^s$

case 1: $\varepsilon^t \geq \underline{\varepsilon}$**,** $t \notin \mathcal{S}, t \leq T(\gamma)$

[10] For any finite set of dates S, $\max S$ denote the largest date in that set.

[11] This is because $v \leq 1 - e^{-rT^*} + e^{-rT^*}\bar{v}_2^{T^*} + O(r \mid \ln r \mid)$, (see Lemma 6) and because by definition of T^*, $1 - e^{-rT^*} = v/2$.

Proposition 2 implies that $\bar{l}_2^t \geq \bar{l}$ and the analysis is identical to that of the simple case (see case 1).[12]

case 2: $\varepsilon^t < \underline{\varepsilon}$ or $T(\gamma) < t \leq \bar{T}(\gamma, \nu)$**.**

Then (by definition of $\bar{T}(\gamma, \nu)$ and ε^t), we have $\varepsilon_2^t \leq \max\{\underline{\varepsilon}, \varepsilon(\nu)\}$. The analysis is again identical to that of the simple case (see case 2), and we obtain $\pi^t(\tau) \leq \frac{\max\{\underline{\varepsilon}, \varepsilon(\nu)\}}{\rho^K}$ (with $\rho = \min\{p, 1-p\}$), which, for τ_ν large enough and r_ν small enough (both determined independently of r) is smaller than, say, $r^{5/4}$ for all $r \leq r_\nu$ and $\tau \geq \tau_\nu$.

case 3: $t \in \mathcal{S}$ and $t \leq T(\gamma)$ (exceptional dates).

Since $\bar{v}_i^t \geq \gamma v \ \forall i$, $\rho^{t,T_0} \geq \gamma v/2$ (see footnote 12), and a fortiori $\varepsilon_1^t \leq 1 - \gamma v/2$. We may thus apply the analysis of case 1 once again, with a bound ε now equal to $1 - \gamma v/2$ instead of $r^{1/2}$. Under event F, the analysis is unchanged. Under event E, the probability $\pi^{t,T_0}(\tau, \bar{l}, \varepsilon \mid \ln r \mid)$ is now comparable to $[r \mid \ln r \mid^2]^{1-\lambda/\tau}$, which is smaller than $r^{4/5}$ for r small enough and τ large enough.

To obtain an upperbound on $\sum_{\substack{t \in \mathcal{S} \\ t \leq T(\gamma)}} \pi^t(\tau)$, we show that there are at most $r^{-3/5}$ exceptional dates before $T(\gamma)$. At any such date, we have $\bar{v}_i^t \geq \gamma v$ (by definition of $T(\gamma)$). So, during any $(\frac{\gamma v}{2r})$ consecutive periods, there cannot be more than $\mid \ln r \mid r^{-1/2}$ dates for which $\varepsilon_i^s > r^{1/2}$ (otherwise we would have $\bar{v}_i^t \leq 1 - e^{-\frac{\gamma v}{2}} + O(r)$, contradicting $\bar{v}_i^t \geq \gamma v$). It follows that $\mathcal{S} \bigcap \{t, t \leq T(\gamma)\}$ contains at most

$$2T_0 \mid \ln r \mid r^{-1/2} \frac{\bar{T}}{\gamma v / 2r}$$

dates. This expression is comparable to $\mid \ln r \mid^3 \ r^{-1/2}$, hence it is smaller than $r^{-3/5}$ for any $r \leq r_\gamma$ if r_γ small enough. And the sum $\sum_{\substack{t \in \mathcal{S} \\ t \leq T(\gamma)}} \pi^t(\tau)$ is thus bounded by $r^{-3/5} r^{4/5} = r^{1/5}$. ■

Part 2 (long-run effect). After date $T^*/2$, party 1 cooperates. So even if at $T^*/2$, party 2's likelihood ratio is much higher under the deviation than the level it would have reached had party 1 not deviated, the gap between these two likelihood ratios will vanish at an exponential rate. Since T^* is comparable to $1/r$, we may choose ν such that by date

[12] Note that for all $t \leq T(\gamma)$, $\rho^{t,T_0} \geq \gamma v/2$ (hence ρ^{t,T_0} bounded away from 0). Indeed, by definition, $\rho^{t,T_0} = \Pr_{\sigma_1, \sigma^C}(\widetilde{t}_1 > t \mid \widetilde{t}_1 > t - T_0)$. Besides for dates $t \leq T(\gamma)$, we have $\bar{v}_2^{t-T_0+1} \geq \gamma v$, and by definition of $\bar{v}_2^{t-T_0+1}$ (see Equation (5)):

$$\bar{v}_2^{t-T_0+1} \leq 1 - e^{-rT_0} + \Pr_{\sigma_1, \sigma^C}\{\widetilde{t}_1 > t \mid \widetilde{t}_1 > t - T_0\}$$

$3T^*/4$, this gap reduces to a small fraction of $\varepsilon(\nu)$ with probability close to 1 (recall that $\varepsilon(\nu) = \exp -\frac{\nu}{r}$).

Having chosen ν as above, we will show that the signals received by player 2 after date $T_1(\nu)$ will bear much more weight on her behavior than the signals she received before $T^*/2$. [The reason is that by definition of $T_1(\nu)$, the probability $\varepsilon_i^{T_1(\nu)}$ is larger than $\varepsilon(\nu)$ for some i, and that without loss of generality, we may assume that $\varepsilon_1^{T_1(\nu)} \geq \varepsilon(\nu)$.[13]] And since player 1 cooperates after $T_1(\nu)$, the effect of the deviation σ^τ after date $T_1(\nu)$ should be small. Proposition 5 formalizes this intuition.

Proposition 5 *Let* $\nu^* = \frac{m \ln(1-v/2)}{20}$, $T_1 = T_1(\nu^*)$, *and* $T_2 = T_1 + \frac{2|\ln r|}{|\ln(1-p)|}$. *There exists* B, $r_0 > 0$ *such that for all* $r \leq r_0$, $\sigma \in \mathcal{E}(v, r)$, *and* $T \geq T_2$,

$$\sum_{T_2 \leq t \leq T} \pi^t(\tau) \leq B \mid \ln r \mid \sum_{t \leq T} \pi^t(\infty) + r. \tag{17}$$

Remark: Note that if in addition we choose r is below r_{μ,ν^*}, then we may also apply Proposition 4 to bound the right hand side of (17). Thus we obtain that if $r \leq \min\{r_0, r_{\mu,\nu^*}\}$, then

$$\sum_{T_2 \leq t \leq T(\mu)} \pi^t(\tau) \leq B \mid \ln r \mid r^{1/5} + r$$

Proof. The result will follow from showing that after T_2 the deviation σ^τ has less effect on party 2's behavior than one bad signal received by party 2 between date T_1 and date T_2. We start with some notation. Consider a date $t > T_2$. We decompose the sequence of signals received by player 2 before date t into four sequences:

$$h_0 = (y_2^s)_{s, s \leq \frac{T^*}{2}}; \qquad x = (y_2^s)_{s, \frac{T^*}{2} < s < T_1}; \qquad z = (y_2^s)_{s, T_1 \leq s < T_2}; \qquad h_1 = (y_2^s)_{;s, T_2 \leq s < t}.$$

We denote by H_0, X, Z and H_1 the set of such sequences, respectively. Let $l_0 = l_2(h_0)$ be the likelihood ratio after the history of signals h_0. At any date $s, T_1 \leq s \leq t$, the likelihood ratio is a function of l_0, x, z and h_1, which we write $L^s[l_0, x, z, h_1]$.

We also denote by $\underline{z}$ the sequence in Z that consists only of bad signals. For any realization $z \neq \underline{z}$, we define the transformed sequence $\phi(z)$ that coincides with z except for the first good signal received, which we replace by a bad signal.

We now define the following three events:

$$\begin{aligned} E &= \{\exists s,\ T_2 \leq s \leq t, L^s[0, x, z, h] \geq \bar{l}_2^s\}; \\ \widehat{E} &= \{z \neq \underline{z}, \text{and } \exists s,\ T_2 \leq s \leq t, L^s[0, x, \phi(z), h_1] \geq \bar{l}_2^s\}; \\ \bar{E} &= \{\exists s,\ T_2 \leq s \leq t, L^s[\bar{L}, x, z, h_1] \geq \bar{l}_2^s\} \end{aligned}$$

[13] In case $\varepsilon_i^{T_1(\nu)} \geq \varepsilon(\nu)$ holds for $i = 2$ only, we may just exchange the role of player 1 and 2 and consider a deviation of player 2.

Step 1: We show that the following inequality holds:

$$\Pr_{\sigma^C,\sigma^C}\{\bar{E}\} \leq \Pr_{\sigma^C,\sigma^C}\{\widehat{E}\} + 2r^2 \tag{18}$$

This is the main step of the proof. Player 1's early defections may drive up player 2's likelihood ratio, and either player 2 switches to defect by date $T^*/2$, or $l_2(h_0) \leq \bar{L}$. In the latter case, inequality (18) says that a single additional bad signal received by player 2 between date T_1 and T_2 has an even stronger impact on player 2's likelihood ratio than whatever occurred before date $T^*/2$, (as long as player 2 did not switch to defect before T_2).

To prove inequality (18), consider $z \neq \underline{z}$ and let t_1 denote the first date $t \geq T_1$ at which party 2 receives a good signal. To facilitate exposition, we write $\bar{L}^t = L^t[\bar{L}, x, z, h]$ and $L^t = L^t[0, x, \phi(z), h]$. Equation (6) (which describes how the likelihood ratio evolves as a function of the history of signals) implies, for all t, $\frac{T^*}{2} \leq t \leq t_1$:

$$\begin{aligned}
\bar{L}^{t+1} - L^{t+1} &\leq [\bar{L}^t - L^t]\frac{\alpha(y_2^t)}{1-\varepsilon_1^{t+1}} \text{ for all } t, \frac{T^*}{2} \leq t \leq t_1 \\
\bar{L}^{t_1+1} - L^{t_1+1} &\leq \frac{\alpha(y_2^{t_1})}{1-\varepsilon_1^{t_1+1}}\{[\bar{L}^{t_1} - L^{t_1}] - \frac{\alpha(\underline{y}) - \alpha(\bar{y})}{\alpha(\bar{y})}L^{t_1}\} \text{ and} \\
L^{t_1} &\geq \prod_{T_1 \leq t < t_1} \frac{\alpha(y_2^t)}{1-\varepsilon_1^{t+1}}\varepsilon(\nu^*)
\end{aligned}$$

Combining this three inequalities implies that, for some strictly positive scalar g, we have:

$$\bar{L}^{t+1} - L^{t+1} \leq g[\bar{L} \prod_{s, T^*/2 \leq s < T_1} \frac{\alpha(y_2^s)}{1-\varepsilon_1^{s+1}} - \frac{\alpha(\underline{y}) - \alpha(\bar{y})}{\alpha(\bar{y})}\varepsilon(\nu^*)] \tag{19}$$

Lemma 3 implies that the product $\prod_{s, T^*/2 \leq s < t_1} \alpha(y_2^s)$ is smaller than $e^{-\frac{|m|T^*/4}{4}}$ with probability $1 - r^4$. It is also easy to check that $\prod_{s, T^*/2 \leq s < T_1}(1-\varepsilon_1^s) \geq \frac{v}{2}(1-\varepsilon(\nu^*))^{\bar{T}}$.[14] The scalar ν^* has been chosen precisely to ensure that

$$\frac{2\bar{L}}{v}(1-\varepsilon(\nu^*))^{\bar{T}} e^{-\frac{|m|T^*/4}{4}} < \frac{\alpha(\underline{y}) - \alpha(\bar{y})}{\alpha(\bar{y})}\varepsilon(\nu^*)$$

for any $r \leq r_0$, with r_0 small enough. The right hand side of (19) is therefore negative with probability $1 - r^4$. It follows that:

$$\Pr_{\sigma^C,\sigma^C}\{L[\bar{L}, x, z, h] \leq L[0, x, \phi(z), h] \mid z \neq \underline{z}\} \geq 1 - r^4. \tag{20}$$

[14]Indeed i) for dates $t \in \{\frac{3T^*}{4}, ..T_1\}$, we have $\varepsilon_1^{s+1} \leq \varepsilon(\nu^*)$; and ii) since v is close to $(1-\varepsilon_1^0)\bar{v}_2^0$ by Lemma 6, since $1 - e^{-rT^*} = v/2$, and since

$$\bar{v}_2^0 \leq 1 - e^{-r\frac{3T^*}{4}} + \Pr_{\sigma_1,\sigma^C}\{\widetilde{t}_1 > \frac{3T^*}{4}\},$$

we must have $\prod_{s \leq \frac{3T^*}{4}}(1-\varepsilon_1^s) \equiv \Pr_{\sigma_1,\sigma^C}\{\widetilde{t}_1 > \frac{3T^*}{4}\} \geq v/2$.

Since $\Pr_{\sigma^C,\sigma^C}\{z=\underline{z}\} \leq r^2$ (because $T_2 = T_1 + \frac{2|\ln r|}{|\ln(1-p)|}$), we obtain the desired result.

Step 2: $\sum_{T_2 \leq s \leq t} \pi^s(\tau) \leq \Pr_{\sigma^C,\sigma^C}\{\bar{E}\}$.

Indeed, for any private history $\widetilde{h}_2^t$ for which the history of signals is given by (h_0, x, z, h_1) and such that $l_2(h_0) < \bar{l}_2^{T^*/2} (\leq \bar{L})$ and $l_2^t(h_2^t) \geq \bar{l}_2^t$, we have $L^t[l_2(h_0), x, z, h_1] \geq \bar{l}_2^t$ by definition of L^t, hence we must also have $L^t[\bar{L}, x, z, h_1] \geq \bar{l}_2^t$.

Step 3: $\Pr_{\sigma^C,\sigma^C}\{\widehat{E}\} \leq (T_2 - T_1)\frac{1-p}{p} \Pr_{\sigma^C,\sigma^C}\{E\} + r^2$.

Note that for any $z_0 \neq \underline{z}$,

$$\Pr_{\sigma^C,\sigma^C}\{\widehat{E} \mid z = z_0\} \equiv \Pr_{\sigma^C,\sigma^C}\{E \mid z = \phi(z_0)\}.$$

For any $z_1 \in Z$, let $Z(z_1) = \{z \in Z, \phi(z) = z_1\}$. We have:

$$\Pr_{\sigma^C,\sigma^C}\{\widehat{E} \mid z \neq \underline{z}\} = \sum_{z_1 \in Z} \Pr_{\sigma^C,\sigma^C}\{E \mid z_1\} \Pr_{\sigma^C,\sigma^C}\{z \in Z(z_1) \mid z \neq \underline{z}\}$$

For any $z \in Z(z_1)$, $\Pr_{\sigma^C,\sigma^C}(z) \leq \frac{1-p}{p} \Pr_{\sigma^C,\sigma^C}(z_1)$. Since the set $Z(z_1)$ contains at most $(T_2 - T_1)$ elements, and since $\Pr_{\sigma^C,\sigma^C}\{z = \underline{z}\} \leq r^2$, we obtain the desired inequality.

Step 4: $\Pr_{\sigma^C,\sigma^C}\{E\} \leq \sum_{s \leq t} \pi^s(\infty)$.

Indeed, if $L^s[0, x, z, h_1] \geq \bar{l}_2^s$, then for any h_0, $L^s[l_2(h_0), x, z, h_1] \geq \bar{l}_2^s$, and player 2's likelihood ratio reaches a threshold $\bar{l}_2^s$ at s or before s (and possibly before T_2, which explains why we sum $\pi^s(\infty)$ over *all* dates before t.)

Combining these steps and recalling that $T_2 - T_1 = \frac{2|\ln r|}{|\ln(1-p)|}$ permit to conclude the proof of Proposition 5. ■

Proof of Proposition 3.

Choose ν^*, T_1 and T_2 as in Proposition 5. Consider first the deviation $\sigma^{\tau_{1/2,\nu^*}}$. This deviation has a negligible effect on party 2's behavior before date $\bar{T}(1/2, \nu^*)$ (by Proposition 4). If there exists a date $t \leq \bar{T}(1/2, \nu^*)$ such that $\bar{v}_1^t < \frac{G(\tau_{1/2,\nu^*})}{2}$, then Proposition 3 holds with $\tau^{**} = \tau_{1/2,\nu^*}$ and $T^{**} = t + 1$.

Otherwise, choose $\gamma = \frac{G(\tau_{1/2,\nu^*})}{4v}$, $\tau^{**} = \tau_{\gamma,\nu^*}$, $\mu = \frac{G(\tau^{**})}{2v}$, $T^{**} = T(\mu) + 1$.

By construction, $\bar{T}(\gamma, \nu^*) > T_2$.[15] Hence the short-run (which ends at $\bar{T}(\gamma, \nu^*)$) and the long run (which starts at T_2) overlap. By choosing $r \leq \min\{r_{\gamma,\nu^*}, r_0, r_{\mu,\nu^*}\}$, we ensure that the deviation $\sigma^{\tau^{**}}$ has a negligible effect on party 2's behavior not only before $\bar{T}(\gamma, \nu^*)$ (thanks to Proposition 4 and because $r \leq r_{\gamma,\nu^*}$), but also between $\bar{T}(\gamma, \nu^*)$ and $T(\mu)$ (thanks to Proposition 5 and 4 and because $r \leq \min(r_0, r_{\mu,\nu^*})$). Since at date $T^{**} = T(\mu)+1$, player

[15]This is because i) for all dates $t \leq \bar{T}(1/2, \nu^*)$, hence for all dates $t \leq T_1$, $\bar{v}_1^t \geq \frac{G(\tau_{1/2})}{2} = 2\gamma v$, and ii) for all dates $t \in \{T_1, ..., T_2\}$, $\bar{v}_1^t$ cannot be smaller than $\bar{v}_1^{T_1}$ by more than $O(r \mid \ln r \mid)$ (because $\bar{v}_1^{T_1} \leq 1 - e^{-r(t-T_1)} + \bar{v}_1^t$).

1's continuation value $\bar{v}_1^{T^{**}}$ is smaller than $\mu v = \frac{G(\tau_\gamma)}{2}$ by definition of $T(\mu)$, Proposition 3 holds with $\tau^{**} = \tau_{\gamma,\nu^*}$ and $T^{**} = T(\mu) + 1$. ■

6 Discussion

Public monitoring vs. private monitoring. A necessary condition for cooperation is that deviations be detected and punished. This further requires that players take into account their past observations, and use them in a way that deters deviations. Learning is a natural candidate for obtaining history dependent strategies:

- If players are uncertain about their opponent's action, they may use their observations to learn about their opponent's past actions.
- If there is some correlation between the opponent's past actions and future play, these observations may thus be used to improve the predictions about the opponent's behavior, and choose an optimal behavior accordingly

When monitoring is private, these two conditions (uncertainty about past actions and correlation between past and future actions) seem to be necessary conditions for obtaining history dependent strategies, as negative results have been found in contexts where either of these conditions fail. In Matshushima (1991), or Bagwell (1995), pure strategies are considered, there is no uncertainty about the opponent's strategy, no learning may occur in equilibrium, and as a result, players do not take into account past observations. In Mailath and Morris (1999) twice repeated convention game,[16] the static game is such that in equilibrium, a player's action cannot depend on payoff irrelevant private information. Hence, even if mixed strategies were played in the first period so that learning about past actions could occur, it would not improve the predictions for the second period. As a result, history-dependent strategies cannot be generated endogenously, and a static Nash equilibrium is played in each period.

In comparison, mixed strategies are key in the positive result of Sekiguchi; and the negative results mentioned above do not carry over once uncertainty or small exogenous correlation between past and future play is introduced: van Damme and Hurkens (1997) have analyzed Bagwell's game where mixed strategies are allowed. And in a perturbed version of Mailath and Morris' game, introducing a small correlation across periods by having types of players who would be playing the cooperative strategy in both periods,

[16]The static convention game is due to Shin and Williamson (1996).

learning about past actions would occur, and the information acquired would be payoff relevant (hence used).

One interpretation of our result is that the two conditions mentioned above are not sufficient to support cooperation. Our analysis suggests the following explanation.

First, equilibrium behavior sets the priors players use to learn from their observations, and this constrains the kind of test on their opponent's behavior that players will perform in equilibrium. Our restriction to trigger strategies imposes that players perform likelihood ratio tests (of the type $\mathbf{l}_i^t[\widetilde{h}_i^t] \leq \bar{l}_i^t$).

This constraint alone however is not sufficient to undermine cooperation. Consider a likelihood ratio test of the following form:

$$\frac{1}{T} \sum_{t-T \leq s < t} \ln \alpha(y_i^s) \leq \gamma, \tag{21}$$

and assume that a player switches to defect when it fails. In the context of public monitoring (where $y_1^s = y_2^s$), if the parameters T and γ are well chosen, such a likelihood ratio test would allow players to support cooperation. Of course, the test would fail with positive probability in any period, and the rate ε at which players switch to defect in equilibrium would be a function of T and γ, say

$$\varepsilon = \varepsilon(T, \gamma). \tag{22}$$

Efficient outcomes would not be possible for reasons explained in Radner Myerson Maskin. Nevertheless, if signals are sufficiently informative, one can find an incentive compatible pair (T, γ) for which ε is comparable to the discount rate r, and values above the mutual minmax would be supported as equilibrium outcomes.

In the context of private monitoring, the likelihood ratio test $\mathbf{l}_i^t[\widetilde{h}_i^t] \leq \bar{l}_i^t$ does not exactly take the same form as the one described above, since all past observations will generally be taken into account. However, the range of dates that effectively contribute to the value of $\mathbf{l}_i^t[\widetilde{h}_i^t]$ depends on the rate at which players switch to defect in equilibrium.[17] If this rate (say ε) is high, the relevant range of dates on which the likelihood ratio test is performed is small. So adjusting the range of dates T on which the likelihood ratio test is performed is possible, but this implicitly determines the rate ε at which players switch to defect, that is,

$$\varepsilon = \varepsilon(T). \tag{23}$$

Yet this is not the only new constraint: incentives to trigger the punishment have to be provided. To see what this implies, observe that any triplet (ε, T, γ) determines each

[17] See Lemma 1 for the dependence of $l_i^t[\widetilde{h}_i^t]$ on ε_j^t.

player's value from continuing to cooperate or switching to defect. Hence it also determines (see Lemma 1) the threshold $\bar{\gamma}(\varepsilon, T, \gamma)$ at which players have the incentives to switch to defect. So providing players with appropriate incentives to trigger a punishment requires that we also set

$$\gamma = \bar{\gamma}(\varepsilon, T, \gamma) \tag{24}$$

Since any test of the form (T, γ) above will induce players to switch to defect at some rate $\varepsilon(T, \gamma)$, Equations (22), (23) and (24) have to be satisfied simultaneously. The problem is that these three equations typically single out the test that may be performed in equilibrium, and yet incentives for players to cooperate have not been checked (only incentives to trigger the punishment have been checked). For the simple case analyzed in Section 5.3. for example, we obtain that ε is at most comparable to $r^{7/6} << r$.[18,19]. In contrast, when monitoring is public, deterring all deviations requires that players agree on a pair (T, γ) such that $\varepsilon(T, \gamma)$ is comparable to r. Thus under private monitoring, the length T of the reviews tend to be too long compared to what it should be to deter short-run deviations. Such deviations are not deterred.

To summarize, two factors contribute to the failure of cooperation when trigger strategies are used. First, players learn from their private observations, taking as given the strategies followed by their opponent. This lead players to likelihood ratio tests of the form (21) under the constraint (23). Second, incentives to effectively punish the opponent have to be provided, which generates the additional constraint (24). Because a player who switches to defect generally gets a very small payoff even when his opponent still cooperates (see Lemma 4), the threshold solving (24) is very large, and players do not react aggressively enough to bad signals.[20]

[18]The general case is complicated by the fact that transition rates are not necessarily stationary over time, but in that case too, on average, transition rate tend to be too small than what they should be to deter short-run deviations.

[19]In our companion paper Compte (1995), defections are assumed to be irreversible, so Equations (23) and (24) are essentially the only two equations that need to be checked. The analysis performed in that paper confirm that the transition rate must be small compared to r.

[20]One interpretation is that defecting for ever is a severe punishment that severely hurts the punisher too. For that reason, he waits for strong evidence before exerting that punishment. The insight that very long punishment makes cooperation fragile appears in the literature on repeated game with a large population (see Kandori (1992) and Ellison (1994)). The idea is that small imperfections would induce few players to switch to defect, and that defections would then spread quickly to the entire population. In our context, cooperation also breaks down eventually, whether monitoring is private or public. However cooperation must break down very fast when monitoring is private.

On the robustness of our result. Our result has been shown to be robust to a class of perturbations of the game. We wish to discuss other directions of robustness that ought to be investigated, and for which we do not have an answer yet.

Conditional independence. A recent paper by Mailath and Morris has investigated the robustness of perfect public equilibria to small changes in the information structure. A similar direction should be taken with our result. We should depart from the pure private monitoring case (conditional independence) and investigate whether a small correlation in the signal structure of the stage game could undermine our result.

Let me briefly mention two reasons why we think that a small correlation in the signal structure may not undermine our result. First, a correlation may actually help the players guess how they are performing in the test carried out by their opponent. So they would have to be provided incentives to cooperate even after receiving positive information concerning their own performance. A result by Compte (1998) and Kandori and Matsushima (1998) actually supports this intuition. Independence is key to their efficiency result (obtained in the two player's case), because it prevents players from guessing what their opponent's signals are, thereby reducing the number of incentive conditions that have to be checked.

Second, repetition may not enable players to create a strong correlation out of a small one. To illustrate why, consider a sequence of signal profile $((y_1^0, y_2^0), ..., (y_1^t, y_2^t), ...)$, where each signal profile $(y_1^t, y_2^t) \in \{0, 1\} \times \{0, 1\}$ is generated by the following distribution:

$y_1 \backslash y_2$	1	0
1	$\frac{1}{2}(1+\rho)$	$\frac{1}{2}(1-\rho)$,
0	$\frac{1}{2}(1-\rho)$	$\frac{1}{2}(1+\rho)$

with ρ small,[21] and let us consider the events

$$A_i^{a,T} = \{\frac{1}{T}\sum y_i^t \leq a\}.$$

We claim that there does not exist a and T such that

$$\Pr\{A_1^{a,T} \mid A_2^{a,T}\}$$

is close to 1. The reason is that either $a < \frac{1-\rho}{2(1+\rho)}$, in which case $\Pr\{A_1^{a,T} \mid (0, ..., 0)\}$ remains very small even when T gets large (hence $\Pr\{A_1^{a,T} \mid A_2^{a,T}\}$ is small too, a fortiori); or $a \in [\frac{1-\rho}{2(1+\rho)}, \frac{1}{2}]$, in which case the event $A_2^{a,T}$ is poorly informative (because ρ is small).

[21] The parameter ρ mesures the correlation between the signals received by each player.

Of course, we should not limit ourselves to this particular class of events. However, the claim above suggests that obtaining $p-$common beliefs events with p close to 1 will be difficult.

On the structure of the strategies used. As mentioned at the start of this Section, uncertainty about past actions and correlation between past and future actions seem to be necessary conditions for supporting cooperation. And the trigger strategies considered in this paper make it easy for players to infer future behavior from past signals.

The restriction to trigger strategies however is clearly a limitation of our results. In particular, our work should be contrasted with that of Sekiguchi (1995), who shows that in a repeated prisoner's dilemma with private but almost perfect monitoring, almost efficient outcomes can be sustained. The construction is as follows. Sekiguchi first shows that for some discount δ_0, there exists an almost efficient equilibrium σ_0 (that equilibrium is actually a trigger strategy equilibrium: players switch to defect after the first bad signal). Then for any δ, T such that $\delta^T = \delta_0$, the original game can be decomposed into T identical and independent *games*: for any $k \in \{1, \ldots, T\}$, the k^{th} *game* is played at date k, $k+T$, $k+2T$, and so on. The actual discount factor for each of those *games* is thus equal to $\delta^T = \delta_0$. Playing σ_0 in each of those games is therefore an equilibrium strategy of the original game, and it yields an almost efficient outcome.[22]

Note that the possibility of decomposing the original game into T independent games is key to the construction above. In some contexts, such a decomposition may not be possible. Here are two examples:

1) There is an initial positive probability that a player is a bad type that always defects. in the equilibrium described above for example, player 1 continues to cooperate with probability close to 1 even if all past $N = r^{-1/2}(< T)$ signals are bad. Yet after such an history, the probability of facing a bad type would increase up to $1 - o(r)$ and player 1 should rather play defect at all later dates.

2) There is a small uncertainty about the date at which signals are received. Then a single occurrence of a bad signal will tend to contaminate the neighboring *games*, and defections will propagate to all *games.*

More generally, whether cooperation can be sustained in perturbed games (in the spirit of the perturbations we considered) is an open question (even when signals are almost perfect or almost public).

[22] Bashkar and van Damme (1997) use a similar technique.

7 Conclusion

Our result sheds some light on why it may be difficult to support cooperation when public signals are absent. Supporting cooperation requires that deviations be detected and punished, which further requires that a player's past observations affect the way he will behave in the future.

When signals are public and players use public strategies, it is easy to ensure that past observations matter: for each history of public signals, we may choose to coordinate future play on a particular continuation equilibrium. We may thus *force players* to take into account *any* given past observations. As a result, many tests on players' past behavior can be performed and carried out in equilibrium.

When observations are private, a player, say player i, may take past observations into account, because they may affect his own belief about how his opponent will play in the future. Yet the link between (player i's) past signals and (player j's) future play cannot anymore be exogenously forced upon. It must be derived privately from player i's own inferences. As a result, there are fewer tests that players will be willing to carry out in equilibrium. We find that trigger strategies fail to support cooperation because players would wait too long to trigger a punishment, and as a result, short run deviations would not be deterred.

Although our result has been show to be robust to some perturbations of the game, we have pointed out in the previous section various other directions in which the robustness of our result should be investigated. This is left for future research.

Appendix A

Lemma 3 and 8 are Corollaries of standard statistical results (See for example Dachuna-Castelle and Duflo (1983, page 171), or Deushel and Stoock (1989, page 3). We start with some notations. For any random variable $\widetilde{z}$ taking values in some finite set $Z \subset \Re$, we define the Legendre transform $\phi(\widetilde{z}, t) = E[\exp(t\widetilde{z})]$ and the Cramer transform

$$h(\widetilde{z}, x) = \sup_{t \geq 0} \{xt - \ln \phi(\widetilde{z}, t)\}.$$

When it exists, we also denote by $t^*(\widetilde{z})$ the scalar that solves $\phi(\widetilde{z}, t^*(\widetilde{z})) = 1$. We have:

Lemma 9 *Let $S_n = \sum_{s=1}^{n} z_s$ denote the sum of n independent random variable distributed as $\widetilde{z}$. For any $x > E\widetilde{z}$, we have $h(\widetilde{z}, x) > 0$ and:*

$$\Pr\{\frac{1}{n} S_n \geq x\} \leq \exp[-nh(\widetilde{z}, x)]. \tag{25}$$

If $E\widetilde{z} < 0$ and $\max Z > 0$, $t^*(\widetilde{z})$ exists and for any $X > 1$, we have:

$$\forall n,\ \ \Pr\{S_n \geq \ln X\} \leq \left[\frac{1}{X}\right]^{t^*(\widetilde{z})}.$$

We will actually prove a slightly stronger result. We consider two independent random variables $\widetilde{z}^0$ an $\widetilde{z}^1$ taking values in some finite set $Z \subset \Re$. For any $\pi \in [0,1]$, we define

$$\phi_\pi(t) = (1-\pi)E[\exp(t\widetilde{z}^0)] + \pi E[\exp(t\widetilde{z}^1)] \tag{26}$$

and

$$h_\pi(x) = \sup_{t\geq 0}\{xt - \ln \phi_\pi(t)\}.$$

We also choose t^*_π such that $\phi_\pi(t^*_\pi) = 1$. We have:

Lemma 10 *For $k \in \{0,1\}$, let $S^k_n = \sum_{s=1}^n z^k_s$ denote the sum of n independent random variable distributed as $\widetilde{z}^k$. Choose $n \geq 1$, $m \geq 0$, and set $\pi = \frac{m}{n+m}$. Let $m_\pi = (1-\pi)E\widetilde{z}^0 + \pi E\widetilde{z}^1$. For any $x > m_\pi$, we have $h_\pi(x) > 0$ and*

$$\Pr\{\frac{S^0_n + S^1_m}{n+m} \geq x\} \leq \exp[-(n+m)h_\pi(x)]. \tag{27}$$

If $m_\pi < 0$ and $\max Z > 0$, there exists t^*_π such that $\phi_\pi(t^*_\pi) = 1$. And for any $X > 1$, we have

$$\Pr\{S^0_n + S^1_m \geq \ln X\} \leq [\frac{1}{X}]^{t^*_\pi} \tag{28}$$

Proof. a) Choose $x > m_\pi$. When t is close to 0, we have $xt - \ln\phi_\pi(t) = t(x-m_\pi) + O(t^2)$, which implies $h_\pi(x) > 0$.

b) Note that $E[\exp(t[S^0_n + S^1_m)] = [\phi_0(t)]^n[\phi_1(t)]^m$. Thus we have:

$$\begin{aligned}\Pr\{\frac{S^0_n + S^1_m}{n+m} &\geq x\} = \Pr\{\exp s(\frac{S^0_n + S^1_m}{n+m} - x) \geq 1\} \leq E\exp s(\frac{S^0_n + S^1_m}{n+m} - x) \\ &\leq \exp -(n+m)[\frac{s}{n+m}x - (1-\pi)\operatorname{Log}\phi_0(\frac{s}{n+m}) - \pi\operatorname{Log}\phi_1(\frac{s}{n+m})]\end{aligned}$$

Since the above inequality holds for any s, we get (27).

c) When t is close to 0, $\phi_\pi(t) = 1 + tm_\pi + O(t^2)$. If $\max Z > 0$, $\phi_\pi(t) > 1$ for t large enough. So if $m_\pi < 0$ and $\max Z > 0$, there exists t^*_π that solves $\phi_\pi(t) = 1$.

d) Inequality (27) implies

$$\Pr\{S^0_n + S^1_m \geq \ln X\} \leq \exp[-(n+m)h_\pi(\frac{\ln X}{n+m})] \tag{29}$$

$$\leq \exp[-\min_{x>0}(\frac{h_\pi(x)}{x})\ln X] \tag{30}$$

To conclude, we check that $\min_x \frac{h_\pi(x)}{x} = t_\pi^*$. Indeed, let $t(x) = \arg\max_t\{tx - \ln\phi_\pi(t)\}$. We have $h'_\pi(x) = t(x)$ and when $\frac{h_\pi(x)}{x}$ is minimum, we must have $\frac{h_\pi(x)}{x} = h'_\pi(x)$. Combining these two equalities yields

$$\frac{h_\pi(x)}{x} = h'_\pi(x) = t(x) = t(x) - \frac{\ln\phi_\pi(t(x))}{x}.$$

It follows that when $\frac{h_\pi(x)}{x}$ is minimum, we must have $\phi_\pi(t(x)) = 1$ (or equivalently, $t(x) = t_\pi^*$), which further implies that $\min_x \frac{h_\pi(x)}{x} = t_\pi^*$. ■

We now prove that Lemma 3 and 8 are Corollaries of Lemma 10.

Proof of Lemma 3: Let $\tilde{y}^0$ (respectively $\tilde{y}^1$) denote the random variable taking values in $Y = \{\bar{y}, \underline{y}\}$ and such that $\Pr\{\tilde{y}^0 = \bar{y}\} = p$ (respectively $\Pr\{\tilde{y}^1 = \bar{y}\} = q$). We define the random variables $\tilde{z}^k = \ln\alpha(\tilde{y}^k)$ for $k = 0, 1$. For any $\tau \geq \tau^*$, we have

$$\Pr_{\sigma^\tau,\sigma^C}\{\prod_{t=1}^{T}\alpha(y_2^t) \geq \exp -\frac{|\, m \,|\, T}{4}\} \leq \max_{n, \frac{T-n}{T} \leq \frac{1}{\tau^*}} \Pr\{\frac{S_n^0 + S_{T-n}^1}{T} \geq -\frac{|\, m \,|}{4}\}.$$

Let $h^* = \inf_{\tau \geq \tau^*} h_{\frac{1}{\tau}}(-\frac{|m|}{4})$. Given the definition of τ^*, $m_{\frac{1}{\tau}} \leq \frac{m}{2}$ for all $\tau \geq \tau^*$, hence $-\frac{|m|}{4}$ is bounded away from $m_{\frac{1}{\tau}}$ for all $\tau \geq \tau^*$. Thus $h^* > 0$. Applying Lemma 10 yields

$$\Pr_{\sigma^\tau,\sigma^C}\{\prod_{t=1}^{T}\alpha(y_2^t) \geq \exp -\frac{|\, m \,|\, T}{4}\} \leq \exp -h^* T,$$

which is smaller than r^4 if $T \geq \frac{4}{h^*} \mid \ln r \mid$. ■

Proof of Lemma 8: Let $l^* = \max\{\varepsilon \mid \ln r \mid, l_0\}$. Define g such that $e^g = 1 + \frac{1}{\alpha(\bar{y})|\ln r|}$. We also define the random variables $\tilde{y}^k$ for $k = 0, 1$, as in the proof of Lemma 3. We may assume that $\bar{l} > l^*$ (otherwise the result is trivially satisfied). For any sequence of signals driving player 2's likelihood ratio up to $\bar{l}$ at date t, there exists a last date $t - t_0$ such that $l_2^{t-t_0} \leq l^*$. For any intermediate date $s \in \{t - t_0, ..., t-1\}$, Equation (6) implies

$$\begin{aligned} l_2^{s+1} &\geq l_2^s \frac{\alpha(y_2^{s+1})}{1 - \varepsilon_2^{s+1}} e^g, \text{ hence} \\ \ln \frac{l_2^t \rho^{t,t_0}}{l_2^{t_0}} &\geq \sum_{s=t_0}^{t-1} (g + \ln\alpha(y_2^{s+1})) \end{aligned}$$

Consider the random variables $\tilde{z}^k = g + \ln\alpha(\tilde{y}^k)$ for $k = 0, 1$. Define $\phi_\pi(t)$ as in Equation (26), and S_n^k and m_π as in Lemma 10. We can choose r_0 small enough so that for any $r \leq r_0$ and $\tau \geq 2\tau^*$, $m_{1/\tau} < 0$ (by definition of τ^* and because g is close to 0 when r is close to 0). By Lemma 10, $t_{1/\pi}^*$ is thus well-defined, and we have:

$$\pi^{t,T}(\tau, \bar{l}, l^0) \leq \max_{\substack{n, \frac{t_0 - n}{t_0} \leq \frac{1}{\tau} \\ t_0 \leq T}} \Pr\{S_n^0 + S_{t_0-n}^1 \geq \ln \frac{\bar{l}\rho^{t,T}}{l^*}\} \leq [\frac{l^*}{\bar{l}\rho^{t,T}}]^{1/t_{1/\tau}^*}$$

When $g = 0$ and $\pi = 0$, $\phi_\pi(1) = E \exp \ln \alpha(\widetilde{y}^0) = E\alpha(\widetilde{y}^0) = 1$ (because $p\frac{q}{p} + (1-p)\frac{1-q}{1-p} = 1$). Hence $t^*_\pi = 1$. It follows that for g and π close to 0, $\frac{1}{t^*_\pi} \geq 1 - \lambda(g + \pi)$ for some $\lambda > 0$, which gives us the desired result.

Appendix B

This appendix is devoted to the proofs of Lemma 1, 4, 5, 6 and 7. Throughout this Appendix, we consider a perturbation ζ of the game. The Lemma in the main text will be proved in the context of the perturbed game. (Note that the perturbation for which $\zeta^t_i = 0$ for all i,t is a possible perturbation, and corresponds to the unperturbed game.)

We first introduce some notation that will allow us to deal with the perturbed game. A realization is now an infinite path of actions, signals and types for each player. We denote by $t^\zeta_i \in \mathcal{N} \cup \{+\infty\}$ the date at which player i becomes a type that always defects. Note that we have:

$$\zeta^t_i = \Pr_\zeta\{t^\zeta_i = t \mid t^\zeta_i \geq t\}.$$

The sequence of *transition probabilities* $(\varepsilon^0_j, \ldots, \varepsilon^t_j, \ldots)$ associated to each trigger strategy σ_j is now defined by

$$\varepsilon^t_j \equiv \Pr_{\sigma_i = \sigma^C, \sigma_j, \zeta}\{\widetilde{t}_j = t \mid \widetilde{t}_j \geq t, t^\zeta_i \geq t\}.$$

A history h^t_i now consists of a history of actions, signals (up to date $t-1$) and type (up to date t). The probability $p^t(h^t_i)$ will now refer to

$$p^t(h^t_i) \equiv \Pr_{\mathbf{a}[h^t_i], \sigma_j, \zeta}\{a^t_j = C \mid h^t_i\},$$

A private history $\widetilde{h}^t_i$ refers to a history for which player i has not switched to defect before t and is still a rational type at date t. The definitions of $\mathbf{l}^t_i[h^t_i]$, $v^t_i(\eta, \sigma_j \mid h^t_i)$ and $\bar{v}^t_i(\eta, \sigma_j)$ still apply (we omit here the reference to the perturbation ζ). Finally, we now define the continuation value $\bar{v}^t_i(\sigma_j)$ as:

$$\bar{v}^t_i(\sigma_j) \equiv E_{\sigma_i = \sigma^C, \sigma_j}[1 - e^{-r(\min\{\widetilde{t}_j, t^\zeta_i\} - t)} \mid \widetilde{t}_j > t, t^\zeta_i > t]. \quad (31)$$

Note that these definitions coincide with the ones given in the main text when $\zeta = 0$, since then we have $t^\zeta_i = +\infty$. Also note that the definitions of $\pi^t(\tau)$ and $\pi^{t,T}(.)$ are kept unchanged. In particular, the proof of Proposition 3 applies to any perturbation.[23]

Proof of Lemma 1.(Existence of thresholds).

[23] Except for the derivation of a lower bound on $\rho^{t,T}$. (see Lemma 12 below).

Consider a trigger strategy equilibrium σ and define $\bar{l}_i^t$ as in Lemma 1. Note that $-v_i(\eta_i, \sigma^D)$ is bounded below by $-b(1-\delta)$, and that $\bar{v}_i^t(\eta_i, \sigma_j) - \bar{v}_i^t(\sigma^D, \sigma_j)$ is bounded above by c. So $\bar{l}_i^t \in [0, \bar{L}]$. Consider a history $\widetilde{h}_i^t$. If $\mathbf{l}_i^t[\widetilde{h}_i^t] > \bar{l}_i^t$, then by construction, for all $\eta \neq \sigma^D$,

$$v_i^t(\eta, \sigma_j \mid \widetilde{h}_i^t) = p^t(\widetilde{h}_i^t)\bar{v}_i^t(\eta, \sigma_j) + (1 - p^t(\widetilde{h}_i^t))v_i^t(\eta, \sigma^D) < p^t(\widetilde{h}_i^t)\bar{v}_i^t(\sigma^D, \sigma_j)$$

and it is thus optimal for player i to switch to defect.

If $\mathbf{l}_i^t[\widetilde{h}_i^t] < \bar{l}_i^t$, then we must have $\bar{l}_i^t > 0$ since $\mathbf{l}_i^t[\widetilde{h}_i^t]$ is non-negative. So by definition of $\bar{l}_i^t$, there must exist a strategy $\eta \neq \sigma^D$ such that $v_i^t(\eta, \sigma_j \mid \widetilde{h}_i^t) > v_i^t(\sigma^D, \sigma_j \mid \widetilde{h}_i^t)$. It follows that after $\widetilde{h}_i^t$, optimal behavior does not consists in defecting for ever. Since σ_i is optimal and since it is a trigger strategy by assumption, it must optimal for player i to continue to cooperate after the history $\widetilde{h}_i^t$. ■

Proof of Lemma 4 (value from defecting).

Choose $\beta > 1$, $T_\beta = \beta \mid \ln r \mid$ and define $\bar{\varepsilon}_i^t = \max_{s \leq T_\beta} \varepsilon_i^{t-s}$. Assume that $\bar{\varepsilon}_1^t \geq \bar{\varepsilon}_2^t$. Since $l_2^{t-s} \geq \varepsilon_1^{t-s}$ for all s, Equation (6) implies

$$l_2^t \geq \bar{\varepsilon}_1^t [\alpha(\underline{y})]^{T_\beta}$$

If party 1 defects from date t until date $t+T$, then from Lemma 3, it must be that

$$l_2^{t+T} \geq l_2^t \exp \frac{T m_D}{2}$$

with probability at least equal to $1 - r^4$. By choosing $T = d_0 \mid \ln r\bar{\varepsilon}_1^t \mid$ and $d_0 = \frac{2}{m_D}[2 + \beta \mid \ln \alpha(\underline{y}) \mid]$, we ensure that l_2^{t+T} is larger than r^{-2} (hence larger than $\bar{L}$ since $\bar{L}$ is comparable to r^{-1}). Player 2 thus switches to defect with probability $1 - r^4$ before date $t+T$, which implies

$$\bar{v}_1^t(\sigma^D, \sigma_2) \leq rT + O((rT)^2)$$

as desired.

Since player 1 actually switches at some date $t - s_0$, $s_0 \leq T_\beta$ with probability $\bar{\varepsilon}_1^t$ (by definition of $\bar{\varepsilon}_1^t$), and since this triggers a response by player 2 before date $t+T$ with probability $1 - r^2$, there must exist a date $t_1 \in \{t - s_0, .., t+T\}$ such that

$$\varepsilon_2^{t_1} \geq \frac{\bar{\varepsilon}_1^t}{T_\beta + T}. \tag{32}$$

If $t_1 \leq t$, then $\bar{\varepsilon}_2^t \geq \varepsilon_2^{t_1}$, and we may apply the analysis above to the case where player 2 starts defecting at t, and obtain that player 1 switches to defect before date $t + T'$ (with $T' = d_0 \mid \ln r\bar{\varepsilon}_2^t \mid$) with probability $1 - r^4$. If $t_1 \geq t$, then we may apply the analysis above to the case where player 2 starts defecting at t_1, and obtain that player 1 switches to defect before date t_1+T' with probability $1-r^4$. In both cases, player 1 switches to defect before date $t+T+T'$. Since (32) implies $T' \leq 2T$ (because $\ln T << T$), we finally obtain:

$$\bar{v}_2^t(\sigma^D, \sigma_2) \leq 3rT + O((rT)^2) \blacksquare$$

Before proving Lemma 5 and 6, we start with the following preliminary result:

Lemma 11 *Choose β as in Lemma 3 and let $\bar{T} = \frac{|\ln r|}{r}$. Also let $\widehat{t} = \min\{\widetilde{t}_2, t_1^\zeta\}$. There exist a, r_0, such that $\forall r \leq r_0$, $\forall \sigma_j$, $\forall T \leq \bar{T}$, we have:*
a) $\Pr_{\eta_\beta^*, \sigma_j, \zeta}\{\widehat{t} \leq \widetilde{t}_1 \leq \widehat{t} + \beta \mid \ln r \mid \quad \mid \widehat{t} \leq T\} \geq 1 - r^2$
b) $\Pr_{\eta_\beta^*, \sigma_j, \zeta}\{\widetilde{t}_1 \geq T \mid \widehat{t} \geq T\} \geq 1 - r^2$

Proof. It will be convenient to let $T_0 = \beta \mid \ln r \mid$. When $\widehat{t} = t$ and $\widetilde{t}_1 < t$, the distribution over histories generated up to date $\widetilde{t}_1$ by the triplet $(\eta_\beta^*, \sigma_j, \zeta)$ coincides with that generated by (σ^C, σ^C). Therefore, for any $t_1 < t$, we have (applying Lemma 3):

$$\Pr_{\eta_\beta^*, \sigma_j, \zeta}\{\widetilde{t}_1 = t_1 \mid \widehat{t} = t\} = \Pr_{\sigma^C, \sigma^C}\{\prod_{s=t_1-T_0}^{t_1} \alpha(y_1^s)\} \geq 1\} \leq r^4,$$

which further implies that for all t,

$$\Pr_{\eta_\beta^*, \sigma_j, \zeta}\{\widetilde{t}_1 \leq t \mid \widehat{t} = t\} \leq tr^4. \tag{33}$$

Similarly, if $\widehat{t} = t$ then either $t_1^\zeta \leq t + T_0$, in which case $\widetilde{t}_1 \leq t_1^\zeta \leq t + T_0$ or $t_1^\zeta > t + T_0$, hence $\widetilde{t}_2 = t$ by definition of $\widehat{t}$. At date $t + T_0 + 1$, under the event $\{\widehat{t} = t, t_1^\zeta > t + T_0\}$, the distribution over signals received by player 1 since t coincides with that generated (σ^C, σ^D). Applying Lemma 3, we obtain:

$$\Pr_{\eta_\beta^*, \sigma_j, \zeta}\{\widetilde{t}_1 > t + T_0 \mid \widehat{t} = t, t_1^\zeta > t + T_0\} \leq \Pr_{\sigma^C, \sigma^D}\{\prod_{s=t}^{t+T_0} \alpha(y_1^s) \leq 1\} \leq r^4. \tag{34}$$

Combining inequalities (33) and (34), and observing that $r^4\bar{T} \leq r^2$ yields the desired result. $\blacksquare$

Proof of Lemma 5: (lower bound on continuation values) Choose $i = 1$, and let $\widehat{t} = \min\{\widetilde{t}_2, t_1^\zeta\}$ and $T_0 = \beta \mid \ln r \mid$. Assume players follow the profile (η_β^*, σ_2) and consider the events $A \equiv \{\widetilde{t}_1 \in \{\widehat{t}, ..., \widehat{t} + T_0\}, \widehat{t} \leq \bar{T}\}$ and $B \equiv \{\widetilde{t}_1 \geq \bar{T}, \widehat{t} \geq \bar{T}\}$. Under event $A \cup B$, the distribution over $\widehat{t}$ coincides with that generated by $(\sigma^C, \sigma_2, \zeta)$, and player 1's expected payoff is equal to $E_{\sigma^C, \sigma_2, \zeta}[1 - e^{-r\widehat{t}} \mid A \cup B] + O(r \mid \ln r \mid)$. Since

$$\Pr_{\eta_\beta^*, \sigma_2, \zeta}\{A \cup B\} \geq 1 - 2r^2$$

by Lemma 11, we get the desired inequalities. ■

Proof of Lemma 6: Consider a trigger equilibrium $\sigma \in \mathcal{E}(r)$ and assume that $\sigma \neq (\sigma^D, \sigma^D)$. (Otherwise the result is immediate). Let $\widehat{t} = \min\{\widetilde{t}_2, t_1^\zeta\}$.

Step 1: We first prove that

$$v_1(\sigma) \leq E_{\sigma^C,\sigma_2,\zeta}[1 - e^{-r\widehat{t}} \mid t_1^\zeta > 0] + O(r \mid \ln r \mid)$$

Let Σ_1 denote the set of trigger strategies for player 1. Now consider the new game where at each date, player 1 would be informed about whether player 2 switches to defect at that date. Let $\bar{\Sigma}_1$ denote the set of trigger strategies for player 1 in that new game. Since $\Sigma_1 \subset \bar{\Sigma}_1$, and since $\sigma_1 \in \Sigma_1$, we must have:

$$v_1(\sigma_1, \sigma_2) \leq \max_{\bar{\sigma}_1 \in \bar{\Sigma}_1} v_1(\bar{\sigma}_1, \sigma_2)$$

Now observe that when player 1 is told that player 2 switches, it is optimal for player 1 to switch to defect too. If player 1 is told that his opponent is still going to cooperate, then we show now that it is optimal for him to continue to cooperate too. Suppose by contradiction that for all $\eta \in \Sigma_1$, $\bar{v}_1^t(\eta, \sigma_2) < \bar{v}_1^t(\sigma^D, \sigma_2)$. Then by equation (4), $v_1^t(\eta, \sigma^D \mid \widetilde{h}_1^t)$ would be strictly negative for all histories $\widetilde{h}_1^t$. This would imply that (in the original game), from date t on, player 1 defects after any history $\widetilde{h}_i^t$. So in the equilibrium σ of the original game, after any $t-$history, both players would defect for ever. By backward induction we would get that both players defect at all dates, contradicting the assumption that $\sigma \neq (\sigma^D, \sigma^D)$.

It follows that when player 1 is allowed to choose a trigger strategy in $\bar{\Sigma}_1$, it is optimal for him to switch to defect exactly at the same date at player 2 does. Let $\bar{\sigma}_1^*$ denote that strategy. We wish to show that

$$v_1(\bar{\sigma}_1^*, \sigma_2) \leq E_{\sigma^C,\sigma_2,\zeta}[1 - e^{-r\widehat{t}} \mid t_1^\zeta > 0] + O(r \mid \ln r \mid).$$

Let $\mathcal{S}_1^\zeta = \{t, \zeta_1^t \geq r^3 \text{ or } t \geq \bar{T}\}$, and consider the event $A = \{t_1^\zeta \in \mathcal{S}_1^\zeta, t_1^\zeta < \widetilde{t}_2\}$. Under event A, player 1 switches to defect before player 2 does, and he may therefore expect to get an expected gain strictly larger than $1 - e^{-r\widehat{t}}$. However, the extra gain cannot be larger than $O(r \mid \ln r \mid)$: when $t_1^\zeta \geq \bar{T}$, the discounted value of the payoffs received after $\bar{T}$ is comparable to r (since $e^{-r\bar{T}} = r$), and when $\zeta_1^t \geq r^3$, player 1 obtains a continuation payoff bounded by $4dr \mid \ln r \mid$ by Lemma 4. Since $\Pr_{\bar{\sigma}_1^*,\sigma_2,\zeta}\{A\} \geq 1 - r^3\bar{T}$ by construction, we get the desired bound on $v_1(\bar{\sigma}_1^*, \sigma_2)$, hence on $v_1(\sigma)$.

Step 2: $v_1(\sigma) = (1 - \varepsilon_2^0)\bar{v}_1^0 + O(r \mid \ln r \mid)$

First observe that $v_1(\sigma) \geq v_1(\eta_\beta^*, \sigma_2)$, and Lemma 5 thus implies $v_1(\sigma) \geq (1 - \varepsilon_2^0)\bar{v}_1^0 - dr \mid \ln r \mid$.

Also, by definition of $\bar{v}_1^0$, we have $E_{\sigma^C,\sigma_2,\zeta}[1-e^{-r\widehat{t}} \mid t_1^\zeta > 0] = \Pr_{\sigma^C,\sigma_2,\zeta}\{\widetilde{t}_2 > 0\}\bar{v}_1^0$. Since the probability that player 2 starts by defecting depends only on (σ_2, ζ), we also have

$$\Pr_{\sigma^C,\sigma_2,\zeta}\{\widetilde{t}_2 > 0\} = \Pr_{\sigma,\zeta}\{\widetilde{t}_2 > 0\} = 1 - \varepsilon_2^0.$$

And step 1 permits to conclude.

Step 3: $\underline{t} = \min\{\widetilde{t}_1, \widetilde{t}_2\}$. We show

$$v_1(\sigma) = E_{\sigma,\zeta}[1 - e^{-r\underline{t}} \mid \underline{t} > 0] + O(r \mid \ln r \mid).$$

Let $\mathcal{S} = \{t, \exists i, \varepsilon_i^t \geq r^3\}$ and consider the events $A = \{\underline{t} \in \mathcal{S}\}$ and $B = \{\underline{t} \geq \bar{T}\}$. Under either event A or event B, player 1 does not obtain a payoff that differs from $1 - e^{-r\underline{t}}$ by more than $O(r \mid \ln r \mid)$.[24] Since $\Pr_{\sigma,\zeta}\{A \cup B\} \geq 1 - r^3\bar{T}$ by construction, we get the desired result.

Final step: The argument of step 3 may be applied to derive $v_2(\sigma)$, which therefore cannot differ $v_1(\sigma)$ by more than $O(r \mid \ln r \mid)$. Combining this observation with step 2 yields the desired result. ■

Proof of Lemma 7.

The strategy $\sigma^{\infty,\beta}$ coincides with η_β^*, and that the latter strategy gives party 1 a payoff close to $v_1(\sigma)$ (from Lemma 5 and 6). We will now compute a lower bound on

$$v_1(\sigma^{\tau,\beta}, \sigma_2) - v_1(\sigma^{\infty,\beta}, \sigma_2)$$

Let $t^\zeta = \min\{t_1^\zeta, t_2^\zeta\}$, and $T_0 = \beta \mid \ln r \mid$. Consider the events $A = \{\forall i, t^\zeta \leq \widetilde{t}_i \leq t^\zeta + T_0$ and $t^\zeta \leq T\}$ and $B = \{\forall i, \widetilde{t}_i \leq t^\zeta$ and $t^\zeta \geq T\}$. Under the assumption of the Lemma, and thanks to Lemma 11, we have:

$$\Pr_{\sigma^{\tau,\beta},\sigma_2,\zeta}\{A \cup B\} \geq 1 - r^{1/6} - 2r^2.$$

In addition, the distribution over t^ζ is independent from the strategies followed by the rational players. Under event $A \cup B$, either both players switch to defect approximately at the same date t^ζ or they switch to defect after T. In either case, the additional payoff that player 1 obtains during the $T^*/2$ first periods from following $\sigma^{\tau,\beta}$ rather than $\sigma^{\infty,\beta}$ is at least equal to

$$(c-1)\Pr_\zeta\{t^\zeta \geq T^*/2 \mid A \cup B\}\frac{1 - e^{-rT^*/2}}{\tau}. \tag{35}$$

[24]Note that Lemma 4 implies that when $\underline{t} = \widetilde{t}_1$, player 1 cannot get a continuation payoff larger than $4dr \mid \ln r \mid$. This also implies that when $\underline{t} = \widetilde{t}_2$, player 1's continuation payoff cannot be smaller than $\frac{-b}{c}4dr \mid \ln r \mid$.

Since $1-e^{-rT^*/2} \geq v/4$ and since $\Pr_\zeta\{t^\zeta \geq T^*/2\} \geq v/2$,[25] the expression above is bounded below by $\frac{(c-1)v^2}{8\tau} + O(r^{1/6})$, hence by $G(\tau)$ for all $r \leq r_0$ if r_0 is small enough.

Since under event B, player 1 loses no more than $\bar{v}_1^T(\sigma_2)$ from following $\sigma^{\tau,\beta}$ instead of $\sigma^{\infty,\beta} = \eta_\beta^*$, we get the desired conclusion. ■

To conclude this Section, we prove the only result used in the proof of Proposition 3 that depends on the perturbation considered.

Lemma 12 *Let* $\rho^{t,T} = \prod_{0\leq s\leq T}(1-\varepsilon_j^{t-s})$. *For any* v, *if* $\bar{v}_i^{t-T} \geq v$ *for all* $s \leq t$ *then* $\rho^{t,T} \geq v - rT$

Proof. Take $i=1$, $j=2$. By definition of $\bar{v}_1^{t-T}$, we have

$$\bar{v}_1^{t-T} \leq 1 - e^{-rT} + \Pr_{\sigma^C,\sigma_2,\zeta}\{\min\{\widetilde{t}_2, t_1^\zeta\} > t \mid \min\{\widetilde{t}_2, t_1^\zeta\} > t-T\}$$

The probability on the right-hand side is equal to $\prod_{0\leq s\leq T}(1-\varepsilon_j^{t-s})$ under the event $\{t_1^\zeta > t\}$ and 0 otherwise, so it is at most equal to $\rho^{t,T}$. The bound on $\rho^{t,T}$ follows. ■

References

[1] Abreu, D., Pearce, D. and Stacchetti, E., "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica,* **58**, 1041–1064 (1990)

[2] Bagwell, K, "Commitment and Observability in Games", *Games and Economic Behavior,* **8**, 271-280 (1995).

[3] Bhaskar, V. and van Damme, E., Moral Hazard and Private Monitoring, mimeo (1997)

[4] Compte, 0. "Communication in Repeated Games with Private Monitoring", *Econometrica*, **66**, 597-626 (1998)

[5] Compte, O., "Sustaining Cooperation without Public Information", 1995, mimeo

[6] Dacunha-Castelle, D. and Duflo, M., *Probabilités et Statistiques*, tome 2, *Problèmes à temps mobile*, Masson, Paris, 1983

[7] van Damme, E. and Hurkens, S., "Games with Imperfectly Observable Commitment", *Games and Economic Behavior*, **21**, 282-308 (1997)

[25] This is because $v \simeq (1-\varepsilon_2^0)v_1^0 \leq (1-e^{-rT^*/2}) + \Pr_\zeta\{t^\zeta \geq T^*/2\}$ and because by definition of T^*, $1-e^{-rT^*} = \frac{v}{2}$.

[8] Deuschel, J.D. and Stook, D., *Large Deviations*, Academic Press, 1989

[9] Ellison, G., “Cooperation in the Prisoner’s Dilemma with Anonymous Random Matching”, *Journal of Economic Theory*, **61**, 567-588 (1994)

[10] Friedman, J. “A Non-cooperative Equilibrium for Supergames”, *Review of Economic Studies*, **38**, 1-12 (1971)

[11] Fudenberg, D. and Levine, D. , “An approximate Folk Theorem with Imperfect Private Information”, *Journal of Economic Theory*, **54**, 26-47 (1991)

[12] Fudenberg, D. and Levine, D. ,“Efficiency and Observability with Long-Run and Short-Run Players”, *Journal of Economic Theory* **62**, 103-135 (1994)

[13] Fudenberg, D., Levine, D. and Maskin, E., “The Folk Theorem with Imperfect Public Information”, *Econometrica* **62**, 997-1040 (1994)

[14] Green, E. and Porter, R., “Noncooperative Collusion under Imperfect Price Formation”, *Econometrica* **52**: 87-100 (1984)

[15] Kandori, M. “Social Norms and Community Enforcement”, *Review of Economic Studies*, **59**, 61-80 (1992).

[16] Kandori, M. and Matsushima, H. , “Private Observations, Communication and Collusion”,*Econometrica*, **66**, 627-652 (1998)

[17] Mailath, G. and Morris, S., “Repeated Games with Almost-Public Monitoring”, mimeo (1999)

[18] Matsushima, H. , “On the Theory of Repeated Games with non-Observable Actions, part I” *Economic Letters*, 35 (1990)

[19] Sekiguchi, T. “Efficiency in Repeated Prisoner’s Dilemma with Private Monitoring”, *Journal of Economic Theory*, **76**, 345-361 (1997)

[20] Shin, H. and Williamson, T. “How much Common Belief is Necessary for a Convention?”, *Games and Economic Behavior,* **13**, 252-268, (1996).

[21] Stigler, G. “A Theory of Oligopoly”, *Journal of Political Economy* 72: 44-61 (1964)