On Sustaining Cooperation without Public Observations*

Olivier Compte†

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

February 2000

*Early versions of this paper had a different title: 'A model of Mutual Trust' in june 1993, and "Endogenous Strategic Uncertainty and Learning from Private Observations" in 1994.

†I am especially grateful to Paul Milgrom for his guidance, and to a referee and the editor for their helpful comments. I thank Nabil Al-Najar, Faruk Gul, Peter Hammond, Philippe Jehiel, David Kreps, Jean Tirole, Jacques Thisse and Dimitri Vayanos for comments and discussions. I also thank seminar participants at UQAM, Canada, séminaire Fourgeaud in Paris, and at ESEM94 in Maastricht. This paper is based on Chapter 1 of my Ph.D. dissertation at Stanford.

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

Abstract: This paper examines a dynamic game in which each player only observes a *private and imperfect* signal on the actions played. Our main result is that in a repeated prisoner's dilemma where defections are irreversible (at least for a long enough period of time), patient enough players may achieve almost efficient outcomes.

Dealing with models of imperfect private monitoring is difficult because i) continuation games are games of incomplete information, hence they do not have the same structure as the original game; In particular, continuation equilibria are correlated equilibria. ii) Players are typically uncertain about their opponents' past observations and actions, and they use their entire own private history to *learn* about these actions; As a result equilibrium strategies are in general non-trivial and increasingly complex functions of past observations.

We bypass these difficulties by looking at correlated equilibria of the original game and find correlated equilibria in which the decision problem faced by each player remains the same over time.

1 Introduction

This paper examines a dynamic game in which each player only observes a private and imperfect signal on the actions played. Our main result is that in a repeated prisoner's dilemma where defections are irreversible (at least for a long enough period of time), patient enough players may achieve almost efficient outcomes.

Recent papers on repeated games with private monitoring have essentially focused on the two following cases: almost perfect monitoring, as in Sekiguchi (1997) or Bhaskar and van Damme (1997)), and almost public monitoring as in Mailath and Morris (1997)).[1] One objective of this paper is to contribute to a better understanding of games where monitoring is *neither close to being perfect, nor close to being public.*

The dynamic game we consider does not have the structure of a standard repeated game since defections have a long term effect on the structure of the continuation game played. Nevertheless most of the difficulties usually encountered when dealing with games where monitoring is private and imperfect will be present in our model, as explained shortly.

Our formal model is a prisoner's dilemma played in continuous time by two players $i = 1, 2$, where we assume that defections are irreversible (players only have to compute the optimal time to switch to Defect). We will show how our analysis extends to cases where defections are not reversible for some period of time (Section 5). We assume that each player receives signals at a rate that depends on the action profile currently played. Throughout our analysis, we assume that signals are good news: signals are more frequent when the opponent cooperates than when he defects.[2] Besides, we assume that as long as both players cooperate, the signals received by each player are independent from one another.

If signals were public, supporting cooperation would be particularly easy. Players could decide to switch *simultaneously* to defect whenever signals become too infrequent. Switching to defect early would be deterred, because it would decrease the frequency of signals and trigger the punishment. Players would also have incentives to follow the punishment, since it is optimal to switch to defect when the other does.

At first glance, finding equilibria supporting cooperation when signals are private might seem to be an easy task too: in order to provide a player, say player 2, with incentives to cooperate, player 1 could switch to defect whenever the signals he receives become too

[1] One exception is Compte (1996), where it is shown that in the context of the repeated prisoner's dilemma, trigger strategies cannot allow players to sustain any level of cooperation above the mutual minimax.

[2] Our analysis can also be carried out if signals are bad news. We refer to the working paper for further details.

infrequent. We would have to check that it is indeed a best response for player 1 to switch to defect when he receives too few signals, but this does not seem to be too arduous either: If player 1 receives too few signals, he may conclude that his opponent is likely to be defecting. Since defections are irreversible, he would then conclude that player 2 is likely to be defecting for ever, hence he should switch to defect now. Indeed, we will show that in equilibrium a player starts to defect when he is sufficiently pessimistic about his opponent being still cooperative.

Working out the details and *finding the statistical test on the frequency of signals that each player ought to perform in equilibrium* turns out to be a difficult task though.

The difficulty is that in equilibrium, we should expect each player to base his decision to switch to defect on the whole history of signals he received. Indeed, *any* signal, whether recent or not, may affect a player's belief about the other's past (hence future) actions, and we should expect this belief to affect the player's decision. How then should each player aggregate past observations in equilibrium?

This complexity is common to all models with imperfect private monitoring. When observations are private and imperfect, each player is uncertain about his opponents' observations, hence about his opponent's past actions.[3] Each player may thus attempt to *learn* about these actions from his own observations, and (to the extent that in equilibrium there is some correlation between past and future actions) a player's belief about his opponent's future behavior will thus be a non trivial function of his own observations. And so will be his strategy.[4]

Another way to assess the difficulty of the problem, is to view each continuation game as a game of incomplete information where each player's history of observations and actions is interpreted as a *type*. Finding an equilibrium is a difficult task because i) the distribution over private observations depends on the actions played, hence (in contrast to standard models of incomplete information) the distribution over types depends on the equilibrium considered - so the particular way in which information is incomplete is *endogenous*, ii) given his own type, a player typically has a non-degenerate belief about his opponent's type; iii) the set of types expends over time, (hence so does the number of potentially distinct continuation strategies) and we should therefore expect equilibrium strategies to

[3]To the extent that behavior is not completely independent from past observations.

[4]Fudenberg and Levine (1991) is an early work on repeated games with imperfect private monitoring that deals with this difficulty by pre-defining dates at which the game restarts. Of course, incentives cannot be provided these dates, so Fudenberg and Levine have to relax the notion of equilibrium they use, and focus on approximate equilibria.

be highly non-stationary and increasingly complex.

We find a way out of this difficulty by looking for correlated equilibria. By doing so, we expect to restore some recursiveness into the problem, since the original game then becomes a game of incomplete information, as continuation games are. And indeed, we show that for appropriate initial correlation devices and appropriate choices of strategies, we can find correlated equilibria where the set of continuation strategies used in equilibrium (as well as the set of equilibrium continuation beliefs) remains the same over time; at any date, after any history of signals, and as long as he cooperates, each player's decision problem remains invariant: Recursiveness can be restored on these paths.[5,6]

Our main result is that when players are patient enough there exist correlated equilibria that give players a payoff as close as we want from the efficient outcome.[7] These equilibria share the following features: i) as long as a player cooperates, his opponent switches to defect at an expected rate that is constant over time (and much smaller than the discount rate, as it turns out); ii) Players use threshold strategies: each player switches to defect as soon as the probability that his opponent has not defected yet goes below a given threshold (this probability thus summarizes in equilibrium the history of his past observations). iii) When a player defects, he increases the rate at which his opponent switches to defect.

Although our result is positive and indicates that sustaining cooperation is possible, we wish to suggest a more cautious interpretation of this result. Indeed, our construction may actually suggest that sustaining cooperation may be quite demanding in terms of the sophistication of the players: either an appropriate correlation device has to be implemented, or equilibrium strategies sustaining cooperation (if such strategies then exist) must be highly non stationary. (See the discussion in Section 6).

The rest of the paper is organized as follows: we present the model and the main result in Section 2. In Section 3, we define the class of threshold strategies that we use to construct our equilibria and explain how we may recover a recursive structure. Section 4 shows that threshold strategies are optimal responses. Section 5 concludes the proof of the main result and considers the extension where defection are reversible after some period of time. We

---

[5]In an earlier version of this work, these equilibria were called Stationary equilibria.

[6]Note that since defections are irreversible, incentives need only be checked on such paths.

[7]Note that in our view, the contribution of the paper lies more in the *existence* statement than the efficiency one. Given existence of equilibria that support (some) cooperation, efficiency should not be surprising: since defections are irreversible, deviations are almost perfectly detectable (statistically) after a lapse of time short compared to $1/r$. In particular, private monitoring does not outperform public monitoring, nor does our result contradicts Radner Myerson Maskin's (1985). (See the discussion section for more details on the interpretation of our main result).

discuss our result in Section 6, and compare our model to the literature and conclude in Section 7.

2 The Model

Payoffs and Action Space. I consider a prisoner's dilemma played in continuous time by two players $i = 1, 2$. Flows of revenues are described by the following bi-matrix:

$1\backslash 2$	C	D
C	$1,1$	b,c
D	c,b	$0,0$

where $c > 1 > 0 > b$. For example, if, during the time interval $[t, t+h)$, player 1 plays C (cooperates), and player 2 plays D (defects), then the revenue for player 1 is bh, and the revenue for player 2 is ch.

We assume that *defecting is irreversible*: if a player has defected once, he cannot get back to cooperation. Formally, at any date t, each player i may be in one of two *physical states* $s_i^t \in \{0,1\}$ In state $s_i^t = 1$, player i may either cooperate or defect. In state $s_i^t = 0$, player i may only defect. Which state player i is in at date $t > 0$ depends on his history of play.[8] We denote by

$$X_i^t \in \{C, \{D_{t'}\}_{t'<t}\}$$

player i's history of play at date t, where $X_i^t = C$ means that player i has not started to Defect before t, while $X_i^t = D_{t'}$ means that player i started to defect at t'. We have:

$$\forall t > 0, \qquad s_i^t = 1 \text{ if } X_i^t = C, \text{ and } s_i^t = 0 \text{ otherwise.}$$

We assume that players discount future payoffs at rate r. Let $u_i(t)$ denote the flow of revenue of player i at date t. Each player i computes the average discounted payoff: $r\int_0^{+\infty} e^{-rt}u_i(t)dt$.

Information structure. Players are assumed to receive signals about the actions played by their opponent. Formally, we denote by t_i^n the date of the n^{th} signal received by player i, and we adopt the convention that $t_i^0 = 0$. A *realization* for player i is a (possibly finite) sequence

$$\omega_i = (t_i^0, t_i^1, \ldots, t_i^n, \ldots) \text{ with } 0 = t_i^0 < t_i^1 < t_i^2 < \ldots < t_i^n < \ldots,$$

[8] At this stage, we do not make any assumption concerning the initial physical state. Note however that our results will be valid even if we impose that initially, $s_i^0 = 1$ (that is, each player may start by cooperating).

where the sequence $\omega_i = (0)$ indicates that no signals are received by player i. We denote by Ω_i the set of possible realizations for player i. We also denote by $\Omega = \Omega_1 \times \Omega_2$ the set of realization profiles $\omega = (\omega_1, \omega_2)$. Given a realization ω_i for player i, we denote by ω_i^t the history of signals up to and including time t. That is, for any sequence $\omega_i = (t_i^0, t_i^1, ..., t_i^n, ...)$ such that $t_i^n \le t < t_i^{n+1}$, we set:

$$\omega_i^t = (t_i^0, t_i^1, ..., t_i^n).$$

We denote by Ω_i^t the set of $t-$histories for player i, and we let $\Omega^t = \Omega_1^t \times \Omega_2^t$.

The distribution over the signals observed depends on the action profile played. Note that since defecting is irreversible, we may focus on the distribution over the signals received by a player that has always cooperated.

We assume that when player i cooperates, he receives signals that follow a Poisson process with arrival rate $\mu > 0$ when player j cooperates, and with arrival rate λ when player j defects.[9] Note since $\mu > 0$, signals have full support over Ω_i^t.[10]

In addition, we assume that when both players cooperates, the signals received by each player follow *independent* processes[11,12] Formally:

For any date t, for any measurable subset $E \subset \Omega_j^t$ and for any history of signal $\omega_i^t \in \Omega_i^t$

$$\Pr_{C,C}\{\omega_j^t \in E \mid \omega_i^t\} = \Pr\{\omega_j^t \in E\}$$

where the subscript (C, C) indicates that we consider the distribution on Ω^t that obtains when both players cooperate (up to date t).

In most of our analysis, we will assume that $\mu > \lambda \ge 0$. In other words, signals are good news: they are more likely when the opponent cooperates than when he defects.[13] We should thus expect that if a player receives no signals during a long amount of time, he should become convinced that his opponent has already switched.

To illustrate this signal structure, consider two firms competing for consumers who arrive according to a Poisson process with arrival rate 2μ, and interpret a signal as the arrival of a

[9] For a Poisson process with arrival rate μ, the distribution over the time of arrival of next signal is independent of the past: $\Pr\{t_i^{n+1} > t + h | t_i^n \le t\} = e^{-\mu h}$.

[10] That is, any measurable subset of Ω_i^t has positive probability.

[11] This contrasts with the public information case where signals would be perfectly correlated.

[12] In our continuous time formulation, players are thus uncertain about their opponent's observations in two ways: players are uncertain about the number of signals received by the other player and about the exact dates when those signals were received.

[13] Our analysis can also be carried out for the case $\lambda > \mu > 0$ (where signals are bad news). We refer to the working paper for further details.

consumer. Assume that under a collusive agreement, consumers split half and half between the two firms so that the arrival rate of consumers to each firm is equal to μ. Assume that when a firm defects, the arrival rate of consumers to the firm who still cooperates is 0. Signals are less frequent when the other defects, hence they are good news ($\lambda = 0$).

Strategies and Nash Equilibria. Since a player who has already defected in the past may only Defect in the future, we need only worry about the private histories for which he did not switch yet to defect; a strategy for player i then indicates for each realization ω_i the time $\tau_i(\omega_i)$ at which player i switches to Defect:

Definition 1 *A strategy for player i is a stopping time $\tau_i : \Omega_i \to \Re^+ \cup \{+\infty\}$ measurable with respect to the history of observations:*

$$\forall t, \forall \omega_i \in \Omega_i, \forall \hat{\omega}_i \in \Omega_i, \quad [\omega_i^t = \hat{\omega}_i^t \text{ and } \tau_i(\omega_i) \leq t] \Rightarrow \tau_i(\omega_i) = \tau_i(\hat{\omega}_i) \tag{1}$$

We denote by $\mathcal{T}_i$ the set of strategies for player i. We also denote by τ^D the strategy that prescribes to Defect at $t = 0$, and by τ^C the strategy that prescribes to cooperate at all dates.

Each strategy profile $\tau = (\tau_1, \tau_2) \in \mathcal{T} \equiv \mathcal{T}_1 \times \mathcal{T}_2$ generates a probability distribution over realizations, which in turn induces a probability distribution over future actions and payoffs. Player i's expected revenue is denoted $v_i(\tau)$ We say that a strategy τ_i is *a best response* to τ_j if τ_i maximizes player i's expected revenue given τ_j. This definition extends to the case where τ_j is a probability distribution over player j's strategies.

A strategy profile $\tau = (\tau_1, \tau_2)$ is a *Nash equilibrium* when τ_i and τ_j are best responses to each other.

Continuation strategies and continuation beliefs. Consider a strategy $\tau_2 \in \mathcal{T}_2$ for player 2. At date t, the private history of player 2 that is relevant for determining her continuation strategy consists of her physical state s_2^t and her history of signals ω_2^t. The *continuation strategy* induced by τ_2 at date t under the *private history* (s_2^t, ω_2^t) is denoted $\tau_2 \mid_{t, s_2^t, \omega_2^t}$. The set of continuation strategies induced by τ_2 at t is denoted $\mathcal{C}^t(\tau_2)$.

Even if player 1 knows with certainty that player 2 follows the strategy τ_2, player 1 will not have observed the signals ω_2^t, nor will he know player 2's current state s_2^t. Player 1 will therefore be uncertain about the continuation strategy followed by player 2 from date t on.[14]

[14] In particular, given any history of private observations, a player's belief about his opponent's history

This is precisely why games with private monitoring are difficult to handle: although the original game is a complete information game, continuation games are games of incomplete information. If one starts with a Nash equilibrium τ, then the continuation strategies form a correlated equilibrium of the continuation game, where the correlation is given by the joint distribution induced by τ at date t over the private histories $(s_1^t, \omega_1^t, s_2^t, \omega_2^t)$. (A definition of correlated equilibrium will be given shortly).

Finding a Nash equilibrium is thus complex for two reasons: 1) the problem does not have a recursive structure (the continuation game is a game of incomplete information whereas the original one is not); 2) over time, the set of history profiles grows, beliefs may involve non-trivial distributions over histories, and checking incentives after any history is difficult in general.

Our framework however is simpler than general dynamic games with private monitoring. There are two reasons for this: i) *Defections are irreversible*; so we need only worry about the decision problem of a player who has always cooperated. ii) *Signals are conditionally independent*; so, at any date t, under the event where player 2 has not switched to defect (that is, under the event $\{s_2^t = 1\}$), and if player 1 has not switched yet to defect, player 1's belief about the continuation strategy followed by player 2 is independent of the $t-$history ω_1^t received by player 1.

Formally, a *belief* of player 1 is defined as a pair $\bar{\beta}_1 = (p_1, \beta_1) \in [0,1] \times \Delta(\mathcal{T}_2)$, where $p_1 = \Pr\{s_2 = 1\}$ and where β_1 is the distribution over player 2's strategies under the event $\{s_2 = 1\}$. The distribution β_1 is called player 1's *conditional belief.* Since when $s_2 = 0$, player 2 has no choice but to defect, a belief $\bar{\beta}_1$ induces a (mixed) strategy for player 2. We will thus sometimes abuse notations and for example let $v_1(\tau_1, \bar{\beta}_1)$ denote the expected revenue obtained by player 1 when he follows τ_1 and player 2 conforms to the mixed strategy induced by $\bar{\beta}_1$.

Assume player 2 follows a strategy τ_2. Then, at any date t, after any history ω_1^t, *and if player* 1 *has not switched yet to defect,* player 1's belief is a pair

$$\bar{\beta}_1^t(\omega_1^t, \tau_2) = (p_1^t(\omega_1^t, \tau_2), \beta_1^t(\tau_2)) \in [0,1] \times \Delta(\mathcal{C}^t(\tau_2))$$

where

$$p_1^t(\omega_1^t, \tau_2) = \Pr^t_{\tau^c, \tau_2}\{s_2^t = 1 \mid \omega_1^t\}$$

of observations at t has full support. This is a crucial difference with the public observations case: When players condition their actions on public information only, each player knows in equilibrium his opponent's continuation strategy in any continuation game.

and where $\beta_1^t(\tau_2)$ is the distribution over player 2's continuation strategies under the event $\{s_2^t = 1\}$. Because of independence, the conditional belief $\beta_1^t(\tau_2)$ does not depend on ω_1^t.[15]

Still, even under our two simplifying assumptions (irreversibility and independence), conditional beliefs may be highly non-stationary:

- The set $\mathcal{C}^t(\tau_2)$ may grow over time, and because of our full support assumption ($\mu > 0$), the conditional belief $\beta_1^t(\tau_2)$ is likely to be a non degenerate distribution over $\mathcal{C}^t(\tau_2)$.
- Even if the set of continuation strategies $\mathcal{C}^t(\tau_2)$ does not expend over time, the conditional belief $\beta_1^t(\tau_2)$ may vary over time.

Therefore, even under irreversibility and independence, the decision problem faced by player 1 may still change overtime and become increasingly complex.

One of our objectives is to show that by looking for correlated equilibria of the game, one may restore some recursiveness into the problem, and that for an appropriate choice of the initial correlation device, and an appropriate choice of the initial strategies, one may ensure that the decision problem faced by each player does not change overtime.

Correlated equilibria. A *type* for player i will consist of a pair $\bar{\theta}_i = (s_i, \theta_i)$ where $s_i \in \{0, 1\}$ indicates the physical state of player i. It will be convenient to denote by $\mathbf{s}_i$ the function that associates to each type $\bar{\theta}_i = (s_i, \theta_i)$ the physical state $\mathbf{s}_i[\bar{\theta}_i] = s_i$ of player i.

In the rest of this paper, θ_i will belong to a subset Θ_i of $\Re$.[16] We let $\bar{\Theta}_i = \{0, 1\} \times \Theta_i$ denote the set of types.

A *type-contingent strategy* $\boldsymbol{\tau}_i$ for player i is a function that specifies for each type $\bar{\theta}_i \in \bar{\Theta}_i$ a strategy $\boldsymbol{\tau}_i[\bar{\theta}_i]$ to be followed by player i *in case* $\mathbf{s}_i[\bar{\theta}_i] = 1$. (In case $\mathbf{s}_i[\bar{\theta}_i] = 0$, player i's strategy is constrained to be equal to τ^D.).

For any distribution over player i's type $f \in \Delta(\bar{\Theta}_i)$, let f^1 denote the conditional distribution over $\bar{\Theta}_i$ induced by f under the event $\{s_i = 1\}$, and let $\boldsymbol{\tau}_i[f^1]$ denote the induced (mixed) strategy. We let

$$\bar{\beta}_{(\mathbf{s}_i, \boldsymbol{\tau}_i)}[f] \equiv (\Pr_f\{\mathbf{s}_i[\bar{\theta}_i] = 1\}, \boldsymbol{\tau}_i[f^1]) \in [0, 1] \times \Delta(\mathcal{T}_i)$$

[15]The notation $\Pr_\tau$ indicates that the distribution over realizations and actions induced by τ is being considered.

[16]The parameter θ_i will summarize all the information that is relevent to determine player i's future behavior in case $s_i = 1$. In equilibrium, on path where player i has not switched yet to defect, it will also summarize all the information that is relevent to player i in assessing player j's (expected) behavior: we will have (in the general case)

$$\theta_i = \Pr[s_j[\bar{\theta}_j] = 1 \mid \bar{\theta}_i = (1, \theta_i)].$$

denote the distribution induced by f over player i's state and strategy.

A *correlation device* is defined as a pair $(\bar{\Theta} = \bar{\Theta}_1 \times \bar{\Theta}_2, \xi)$ where ξ is a distribution over $\bar{\Theta}$. For any type $\bar{\theta}_i \in \bar{\Theta}_i$, we let $\xi \mid_{\bar{\theta}_i} \in \Delta(\bar{\Theta}_j)$ denote the belief of player i about the type of his opponent induced by ξ.

A *correlated equilibrium* is a triplet $(\bar{\Theta}, \xi, \boldsymbol{\tau})$ where $\boldsymbol{\tau} = (\boldsymbol{\tau}_1, \boldsymbol{\tau}_2)$ and such that for any $\bar{\theta}_i \in \bar{\Theta}_i$ for which $\mathbf{s}_i[\bar{\theta}_i] = 1$, the strategy $\boldsymbol{\tau}_i[\bar{\theta}_i]$ is a best response against $\bar{\beta}_{(\mathbf{s}_j, \boldsymbol{\tau}_j)}[\xi \mid_{\bar{\theta}_i}]$. We denote that by $v_i(\boldsymbol{\tau}, \xi)$ the expected payoff obtained by player i when both players conform to $\boldsymbol{\tau}$.

Main result

Theorem 1 *For any given μ, λ, b, c with $\mu > \lambda \geq 0$ and $c > 1 > 0 > b$, and for any $\varepsilon > 0$, there exists r_0 such that for any r, $0 < r \leq r_0$, there exists a correlated equilibrium $(\bar{\Theta}, \xi, \boldsymbol{\tau})$ such that $v_i(\boldsymbol{\tau}, \xi) \geq 1 - \varepsilon$ for each player $i = 1, 2$.*

Our main result thus states that we may construct correlated equilibria involving cooperation and that when players are patient enough, these equilibria can yield an outcome as close as we want from the efficient outcome. Note that the statement of our main result implies that the probability $\Pr_\xi[s_i = 0]$ that player i is in physical state 0 at the start of the game is smaller than ε. We will actually prove a stronger result in Section 5, showing that the correlated equilibrium can be chosen so that $\Pr_\xi[s_i = 0] = 0$.

3 Recovering a recursive structure.

As explained in the preceding section, the main difficulty with games with private monitoring is that the decision problem faced by each player may change over time and become increasingly complex. The purpose of this section is to show how (contingent) strategies and initial beliefs may be chosen to ensure that the decision problem faced by each player does not change over time.

3.1 A simple introductory case: $\mu > \lambda = 0$.

The case $\lambda = 0$ is simpler because receiving a signal at date t indicates with certainty that one's opponent has not switched yet to defect, that is, that his physical state is 1. Natural candidates for equilibrium behavior in this context are strategies where the decision to switch to defect only depends on *the lapse of time since the last signal (if any) occurred.* We start by defining strategies that have this property.

For any history ω_i^t, we let $\bar{t}_i(\omega_i^t, t)$ denote the date of the last signal before date t received by player i, and set $\bar{t}_i(\omega_i^t, t) = 0$ in case $\omega_i^t = (0)$. For any $T > 0$, we define the strategy τ_i^T for player i as follows: player i cooperates as long as the lapse of time since the last signal is smaller than T, that is, as long as

$$t - \bar{t}_i(\omega_i^t, t) < T.$$

We consider the set $\mathcal{C}_i^T = \cup_{t \geq 0} \mathcal{C}^t(\tau_i^T)$ of continuation strategies induced by τ_i^T. At any date t, either player i has already switched to defect (that is, $s_i^t = 0$) and his continuation strategy is τ^D, or player i may still start by cooperating at t (that is, $s_i^t = 1$) and his continuation strategy depends only on the lapse of time

$$\theta_i^t = t - \bar{t}_i(\omega_i^t, t)$$

since the last signal was received: he continues to cooperate if $\theta_i^t < T$, and he switches to defect otherwise. We denote by $\boldsymbol{\tau}_i^T[\theta_i^t]$ the corresponding strategy. Let $\Theta_i = [0, +\infty)$. The set of continuation strategies induced by τ_i^T is

$$\mathcal{C}_i^T = \{\boldsymbol{\tau}_i^T[\theta_i], \theta_i \in \Theta_i\}.$$

Although we considered the continuation strategies induced by τ_i^T, it should be clear that the set of continuation strategies induced by $\boldsymbol{\tau}_i^T[\theta_i]$ is also contained in $\mathcal{C}_i^T$. Thus for any initial belief $\bar{\beta}_1^0 = (p_1^0, \beta_1^0)$ of player 1 where β_1^0 has a support included in $\mathcal{C}_2^T$, player 1's belief over player 2's continuation strategy at date t is a pair $\bar{\beta}_1^t = (p_1^t, \beta_1^t)$ where β_1^t also has a support included in $\mathcal{C}_2^T$.

In what follows, we restrict our attention to conditional beliefs that are distributions over $\mathcal{C}_i^T$ for some T: a conditional belief for player 1 will be described as pair

$$\beta_1 = (\mathcal{C}_2^T, B_1(.)) \text{ where } B_1(\theta) = \Pr\{\theta_2 \leq \theta\}.$$

For any initial conditional belief $\beta_1^0 = (\mathcal{C}_2^T, B_1^0)$, player 1's conditional belief at date t is thus a pair $\beta_1^t = (\mathcal{C}_2^T, B_1^t)$. The following Proposition states that for an appropriately chosen initial distribution B_1^0, player 1's conditional belief remains invariant over time.

Proposition 1 *For any $T > 0$, define K such that $\frac{K}{\mu} = e^{-(\mu - K)T}$ and*

$$B_T(\theta) = \frac{1 - e^{-(\mu-K)\theta}}{1 - \frac{K}{\mu}}.$$

Let $\beta_T = (\mathcal{C}_2^T, B_T)$. If $\beta_1^0 = \beta_T$, then at any date t, after any history ω_1^t, and if player 1 has not switched yet to defect, then $\beta_1^t = \beta_T$.

Proof. Assume $B_1^t(\theta) = B_T(\theta)$. During the time interval $[t, t+dt]$, the flow of exit from $[0, \theta]$ is $[B_T(\theta) - B_T(\theta - dt)]$, and the flow of entry is $[1 - B(\theta)]\mu dt$. The overall flow of exit is thus equal to $B_T'(\theta) - (1 - B_T(\theta))\mu$, which is (given the definition of B_T) equal to $KB_T(\theta)dt$. Since Kdt is independent of θ, the conditional distribution remains invariant over time and equal B_T. ■

Note that we already knew from the analysis of Section 2 that thanks to independence, conditional beliefs would not depend on the particular history ω_1^t. The only addition of Proposition 1 is that conditional beliefs may remain invariant over time (under appropriate initial conditions and as long as player 1 cooperates): for any initial belief $\bar{\beta}_1^0 = (p_1^0, \beta_T)$, the dynamics of player 1's belief is fully described by a single parameter, which is, the probability

$$p_1^t(\omega_1^t, \bar{\beta}_1^0) = \Pr{}^t_{\mathcal{T}^C, \bar{\beta}_1^0}\{s_2^t = 1 \mid \omega_1^t\}$$

that player 2 has not switched yet to defect.

At any date, player 1 thus faces the same family of decision problems: given his belief (p, β_T), where $p \in [0, 1]$, should he continue to cooperate or switch to defect?

3.2 The general case. ($\mu > \lambda > 0$).

In this section, our objective is the same as in the previous one: derive an invariant conditional belief $\beta \in \Delta(\mathcal{T}_j)$ over player j's strategies, ensuring that the decision to continue to cooperate or to switch to defect faced by player $i \neq j$ only depends, at any date, on the current probability that player j still cooperates.

The first difficulty is the choice of the strategies in the support of β.

Threshold strategies. In the previous section, we chose strategies such that player 2 would continue to cooperate unless no signals are received during a lapse of time equal to T. The complication with the general case is that we cannot expect equilibrium strategies to have such a simple structure: after receiving a signal, player i will not be certain that his opponent has not switched yet to defect; and his belief about his opponent's physical state will not depend only on the lapse of time since the last signal received, but on the whole sequence of signals.

How will players aggregate information about past observations in this context? In equilibrium, each player, say player 1, can compute the probability p_1^t that his opponent still cooperates, and a natural candidate for equilibrium behavior is that player 1 switches to defect as soon as p_1^t gets below some threshold $\underline{p}$.

Formally, let $\bar{\beta}_1^0 = (p_1^0, \beta_1^0)$ be the initial belief of player 1. For any date $t \geq 0$, and any history of signals ω_1^t, we define the probability

$$p_1^t(\omega_1^t, \bar{\beta}_1^0) = \Pr^t_{\tau^C, \bar{\beta}_1^0}\{s_2^t = 1 \mid \omega_1^t\}$$

that player 2 did not switch to defect before t given that player 1 received ω_1^t, when player 1 always cooperates and player 2 follows $\bar{\beta}_1^0$. The proposed strategy for player 1 is that player 1 continues to cooperate as long as:

$$p_1^t(\omega_1^t, \bar{\beta}_1^0) > \underline{p}$$

and switch to defect otherwise.

The difficulty of course is that we cannot perform any computation and derive further player 1's strategy before we make some assumption concerning the belief $\bar{\beta}_1^0$, hence about player 2's equilibrium behavior.

We will assume (and check later whether this assumption is satisfied) that *as long as player* 1 *cooperates*, player 2 on average switches to defect at a constant rate.[17] Formally we define:

Definition 2 *The (mixed) strategy $\beta \in \Delta(\mathcal{T}_2)$ has the constant switching rate property if there exists a constant K such that for any $t \geq 0$ and $h \geq 0$, we have:*

$$\Pr^t_{\tau^C, \beta}[s_2^{t+h} = 1 \mid \omega_1^t] = e^{-Kh} \Pr^t_{\tau^C, \beta}[s_2^t = 1 \mid \omega_1^t] \tag{2}$$

Player 2 is then said to switch to defect at a constant rate K.

The next proposition shows that the rate K at which player 2 switches to defect fully determines the evolution of player 1's beliefs as a function of the signals received (as long as player 1 cooperates).

Let $\bar{p} \equiv 1 + \frac{K}{\mu - \lambda}$ and define the functions $\phi_s(p)$ and $\phi_{n,K}(p, h)$ as follows:

$$\phi_s(p) \equiv \frac{p\mu}{p\mu + (1-p)\lambda} \tag{3}$$

$$\frac{\bar{p}}{\phi_{n,K}(p,h)} - 1 \equiv (\frac{\bar{p}}{p} - 1)e^{-(\lambda - \mu - K)h} \tag{4}$$

The following Proposition is a straightforward application of Bayes Law:

[17] The motivation for restricting our attention to strategies that have the constant switching rate property is as follows: ultimately, we wish to obtain an equilibrium where player 1's conditional belief remains invariant over time (as long as player 1 cooperates). In any such equilibrium, the conditional distribution over player 2's continuation strategy -conditional on player 2 still being in physical state 1- remains the same over time (as long as player 1 cooperates) by definition. So the rate at which player 2 defects must be constant (as long as player 1 cooperates).

Proposition 2 *Let $\bar{\beta}_1^0 = (p_1^0, \beta_1^0)$ be the initial belief of player 1, and assume that β_1^0 has the constant switching rate property with rate K. Each realization $\omega_1 = (t_1^0, ..., t_1^n, ...)$ defines a (right continuous) path of probabilities $q^{\omega_1, p_1^0}(.)$ starting from $q^{\omega_1, p_1^0}(0) \equiv p_1^0$ and defined by $q^{\omega_1, p_1^0}(t) \equiv p_1^t(\omega_1^t, \bar{\beta}_1^0)$. This path depends only on K and the initial probability p_1^0. Besides if no signals occur during $(t, t+h]$, then*

$$q^{\omega_1, p_1^0}(t+h) = \phi_{n,K}(q^{\omega_1, p_1^0}(t), h),$$

and if a signal occurs at $t > 0$, then

$$q^{\omega_1, p_1^0}(t) = \phi_s(q^{\omega_1, p_1^0}(t^-)).$$

Proof. See Appendix F ■.

Figures 1 gives an illustration of a path $q^{\omega_1, p_1^0}(t)$: it decreases smoothly when no signals occur (because the absence of signal is bad news) and it jumps upward when a signal occurs.[18,19,20]

We are now equipped to define the strategies that we will use to construct our equilibria.

Definition 3 *(The threshold strategy $\tau^{K,\underline{p}}$) Choose $K > 0$ and $\underline{p} < 1$. For any realization ω_i, consider the path $q^{\omega_i, 1}(.)$ (starting from $q^{\omega_i,1}(0) = 1$) as defined in Proposition 2. We let $\tau_i^{K,\underline{p}}(\omega_i)$ be the first date t at which $q^{\omega_i,1}(t) \leq \underline{p}$. That is, under the strategy $\tau_i^{K,\underline{p}}$, a player cooperates as long as $q^{\omega_i,1}(t) > \underline{p}$.*

Clearly, the definition can be extended to the case where the path considered is $q^{\omega_i, \theta_i}(.)$ (which starts from $\theta_i \in [0,1]$ instead of 1). We denote by $\boldsymbol{\tau}_i^{K,\underline{p}}[\theta_i]$ the corresponding strategy, and let

$$\mathcal{C}_i^{K,\underline{p}} \equiv \{\boldsymbol{\tau}_i^{K,\underline{p}}[\theta_i], \theta_i \in [0,1]\}.$$

[18] In the bad news case ($\lambda > \mu$), the same equations would describe the dynamics of beliefs: the path $q^{\omega_1, p_1^0}(.)$ would tend smoothly toward $\overline{p}$ as long as no signal occurs. At $\overline{p} = 1 + \frac{K}{\mu - \lambda}$ (note that $\overline{p} < 1$ in this case), the absence of signals (which is good news) would exactly compensates the fact that player 2 switches to Defect at rate K. When a signal occurs, the path jumps downward.

[19] To illustrate further how one can use Proposition 2 to compute a particular path, consider for example the case where $\omega_1 = (0, t_1^1)$, and let $p_1^0 = p_1^0(\omega_1^0, \beta^0)$ For any $t < t_1^1$,

$$q^{\omega_1, p_1^0}(t) = \phi_{n,K}(p_1^0, t)$$

and for any $t \geq t_1^1$,

$$q^{\omega_1, p_1^0}(t) = \phi_{n,K}\left[\phi_s[\phi_{n,K}(p_1^0, t_1)], t - t_1\right]$$

[20] When $\lambda = 0$ as in the introductory example, $\phi_s(p) = 1$ for all $p > 0$, thus $q^{\omega_i,1}(t)$ depends only on the lapse of time since the last signal was received.

As in the introductory example, if player 1's initial conditional belief has a support included in $\mathcal{C}_2^{K,\underline{p}}$, then at any later date, player 1's conditional belief also has a support included in $\mathcal{C}_2^{K,\underline{p}}$. In what follows, we restrict our attention to conditional beliefs which have their support included in $\mathcal{C}_2^{K,\underline{p}}$ for some $(K, \underline{p})$, and we describe any such conditional belief as a pair[21]

$$\beta_1 = (\mathcal{C}_2^{K,\underline{p}}, B_1) \text{ where } B_1(\theta) = \Pr\{\theta_2 \geq \theta\}.$$

Any initial conditional belief $\beta_1^0 = (\mathcal{C}_2^{K,\underline{p}}, B_1^0)$ induces at date t a conditional belief $\beta_1^t = (\mathcal{C}_2^{K,\underline{p}}, B_1^t)$.

Invariant conditional belief consistent with K. As in the introductory example, our aim is to find an invariant conditional belief β_1^0. This will ensure that player 1's decision problem does not change over time: only the level of p_1 (and not the date) will then affect his decision to switch to defect.

An important difference with the introductory example however is the following: player 1's belief bears on strategies in $\mathcal{C}_2^{K,\underline{p}}$; these strategies for player 2 implicitly assume that player 1 switches to defect at a constant rate K. Our aim is to construct symmetric equilibria where each player switches to defect at the same constant rate K (as long as his opponent cooperates). So we need to check that the invariant conditional belief β_1^0 is consistent with player 2 switching to defect at a constant rate K.

The next proposition states that for an adequate choice of the threshold $\underline{p}$, it is possible to find such a consistent and invariant conditional belief.

Proposition 3 *There exists $K_0 < \mu$ such that for any $K \leq K_0$, there exists a threshold $\underline{p}$ and a distribution B having the following properties:* **i) invariance***: if $\beta_1^0 = (\mathcal{C}_2^{K,\underline{p}}, B)$, then at any date t, after any history ω_1^t, and if player 1 has not switched yet to defect, $\beta_1^t = \beta_1^0$.* **ii) $K-$consistency:** *under $\beta_1^0 = (\mathcal{C}_2^{K,\underline{p}}, B)$, player 2 switches to defect at a constant rate K (as long as player 1 cooperates).*

The proof of this Proposition is in Appendix A. To fix ideas, we show next that $K-$consistent and invariant beliefs can be found by solving a differential equation with appropriate boundary conditions.

[21] Note that the sign of the inequality in the definition of B differs from that given in the introductory case. Also note that the parameter θ_i has a different interpretation. It corresponds here to the probability that one's opponent is in physical state 1 (see also footnote 16), whereas in the introductory example, it corresponded to the lapse of time since the last signal was received.

Consider the following differential equation:

$$-\mu[B(\phi_s^{-1}(\theta)) - B(\theta)] - (\mu - \lambda)\theta(\bar{p} - \theta)B'(\theta) = KB(\theta), \tag{5}$$

along with the boundary conditions:

$$B(1) = 0 \text{ and } B(\theta) = 1 \text{ for all } \theta \leq \underline{p} \tag{6}$$

We show the following proposition:

Proposition 4 *Fix a transition rate $K \in (0, \mu)$, and assume that there exists $(\underline{p}, B)$ that solves the differential equation (5) with boundary conditions (6). Then the conditional belief $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$ is an invariant condition belief for player 1 under which player 2 switches to defect at a constant rate K.*

Proof. Assume $(\underline{p}, B)$ solves the differential equation (5) with boundary conditions (6), and assume that $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$ describes player 2's behavior. Recall that $B(\theta) = \Pr\{\theta_2 \geq \theta\}$. The term $\mu dt[B(\phi_s^{-1}(\theta)) - B(\theta)]$ corresponds to the probability that θ_2 enters the interval $[\theta, 1]$ during the time interval $[t, t+dt]$,[22] while $-(\mu - \lambda)\theta(\bar{p} - \theta)B'(\theta)dt$ may be interpreted as the probability that θ_2 exits from the interval $[\theta, 1]$ during the time interval $[t, t+dt]$.[23] Overall, the flow of exit from $[\theta, 1]$ is thus equal to $KB(\theta)dt$, which implies the desired properties. ■

To conclude this Section, we relate the threshold $\underline{p}$ to the transition rate K when K small.

Proposition 5 *There exist K_0, d_1 and d_2 such that for any $K \in (0, K_0)$ and $(\underline{p}, B)$ solving the differential equation (5) with boundary conditions (6),*

$$\underline{p} \geq \frac{d_1}{|\ln K|} \text{ and } B(1 - K^{1/2}) \geq 1 - d_2 K^{1/2}$$

The proof is in Appendix B. Proposition 5 shows that the distribution B puts most weight on values of θ close to 1 when K is close to 0. The intuition is that players on average learn the truth: when player 1 cooperates, player 2 tends to learn it, and most realizations of her signals will push her beliefs towards $p = 1$. Of course, bad realizations (very infrequent signals) are possible even when player 1 cooperates, and this is why there is always a positive probability that player 2 switches to defect. Yet, as the threshold $\underline{p}$

[22] Entry occurs when a signal is received during dt and $\theta_2 \in [\phi_s^{-1}(\theta), \theta)$ (first approximation).

[23] Exit occurs when no signal is received during dt and when $\theta_2 \in [\theta, \theta')$, with $\phi_{n,K}(\theta', dt) = \theta$.

decreases, the probability that bad realizations drive player 2's belief below $\underline{p}$ decreases at an exponential rate, as it turns out, explaining the relationship between $\underline{p}$ and K.[24]

4 Best Responses

We now turn to best responses. We wish to check that our initial assumption that players would follow threshold strategies was justified. We have:

Proposition 6 *Assume that $s_1^0 = 1$ and that player 1's belief is $\bar{\beta}_1^0 = (p_1^0, \beta)$ where $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$ is invariant and consistent with K. Then there exists a threshold $\underline{q}$ such that it is optimal for player 1 to follow the threshold strategy $\boldsymbol{\tau}_1^{K,\underline{q}}[p_1^0]$.*

The intuition is the same as for our companion paper Compte (1996). At any date t, after any history, only two events matter for player 1: either player 2 has already switched to defect (in which case player 2 follows τ^D), or player 2 has not switched to defect, and player 2 is following β. Player 1's optimal decision depends on the relative weight of these two events. In addition, since β is invariant, player 1's decision problem remains the same over time, and thus the threshold under which player 1 switches to defect also remains the same over time.

Proof. Assume player 1 follows the strategy η. We have:

$$v_1(\eta, \bar{\beta}_1^0) = p_1^0 v_1(\eta, \beta) + (1 - p_1^0) v_1(\eta, \tau^D)$$

If player 1 starts by defecting, he obtains

$$v_1(\tau^D, \bar{\beta}_1^0) = p_1^0 v_1(\tau^D, \beta)$$

Note that for any $\eta \neq \tau^D$, $v_1(\eta, \tau^D) < 0$. So if $v_1(\eta, \beta) \leq v_1(\tau^D, \beta)$ for all η, it is optimal for player 1 to start by defecting, and we set $\underline{q} = 1$. Otherwise, there exists $\eta (\neq \tau^D)$ such that $v_1(\eta, \beta) > v_1(\tau^D, \beta)$ and we may define:

$$\underline{q} = \min\{1, \inf_{\{\eta, v_1(\eta,\beta) > v_1(\tau^D,\beta)\}} \frac{-v_1(\eta, \tau^D)}{v_1(\eta, \beta) - v_1(\tau^D, \beta) - v_1(\eta, \tau^D)}\}$$

For any $p_1^0 > \underline{q}$, it is optimal for player 1 to start by cooperating.

At any later date, player 1's conditional belief remains equal to β. He thus faces the same decision problem and it is optimal for him to stick to the same decision rule, that

[24] Note that for the introductory example, a tighter bound can be obtained, as it is easy to check directly that $\underline{p}$ tends to $1/2$ as K tends to 0 (that is, as T gets very large).

is, to continue to cooperate as long as $\Pr^t_{\tau^C,(p_1^0,\beta)}\{s_2^t = 1 \mid \omega_1^t\} > \underline{q}$. Since β has the constant switching rate property with rate K, this amounts to following the threshold strategy $\boldsymbol{\tau}_1^{K,\underline{q}}[p_1^0]$ (by definition of $\boldsymbol{\tau}_1^{K,\underline{q}}$). ■

To conclude this Section we exhibit various bounds on the payoff obtained by player 1 when he faces a player that plays according to some invariant and K-consistent strategy, as well as an upper bound on the optimal threshold $\underline{q}$.

Proposition 7 *There exist a constant d and r_0 such that for any discount $r \leq r_0$ and $(K, \underline{p}, B)$ satisfying: i) $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$ is invariant and consistent with K, and ii) $K \leq r$, the following holds:*
1) $v_1(\tau^D, \beta) \leq dr \mid \ln K \mid$
2) There exists η^* such that $v_1(\eta^*, \beta) \geq \frac{r}{r+K} - dr \mid \ln r \mid$ and $v_1(\eta^*, \tau^D) \geq -dr \mid \ln r \mid$.
3) Player 1's optimal threshold $\underline{q}$ satisfies $\underline{q} \leq dr \mid \ln K \mid$

The detailed proof is in Appendix C. The argument is very similar to that in Section 4 in Compte (1996).

The first inequality is proved by showing that if player 1 defects during a lapse of time comparable to $\mid \ln K \mid$, player 2 switches to defect with probability equal to $1 - K$ at least.

By switching to defect exactly at the same date as player 2 does, player 1 would obtain $\frac{r}{r+K}$ in the event player 2 follows β, and 0 in the event player 2 follows τ^D. The second property thus says that player 1 can always secure a payoff close to the one he would get if he switched to defect exactly at the same date as player 2 does. The intuition is that for d large enough, it is easy for player 1 to discriminate efficiently between the event "player 2 has switched to defect before date $t - d \mid \ln r \mid$" and the event "player 2 is still cooperating".[25]

Finally, the third property is an immediate consequence of the first two.

5 Main result and Extensions.

This section concludes the proof of our main result and derives two extensions.

Before concluding the proof of our main result, we show that there exists a conditional belief $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$ that is invariant, consistent with K, and for which it is optimal for player 1 to follow the threshold strategy $\boldsymbol{\tau}_1^{K,\underline{p}}$ (that is, with a threshold precisely equal to $\underline{p}$).

[25] A strategy η^* that achieves this is one where player 1 makes reviews every $d \mid \ln r \mid$ units of time and check whether the average number of signals received per unit of time is smaller or larger than $\frac{\lambda+\mu}{2}$.

Proposition 8 *There exists r_0 such that for all $r \leq r_0$, there exist $K \leq r$, $\underline{p} < 1$ and B such that*

i) *the conditional belief $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$ is invariant and consistent with K.*

ii) *if $s_1^0 = 1$ and $\bar{\beta}_1^0 = (p_1^0, \beta)$, then it is optimal for player 1 to follow $\boldsymbol{\tau}_1^{K,\underline{p}}[p_1^0]$.*

The detailed proof of this Proposition is in Appendices B, C and D. We provide here the main idea of the argument.[26,27]

Step 1: We construct a continuous mapping $(K, \underline{p}) \to B_{K,\underline{p}}$ where $B_{K,\underline{p}}$ has the following properties: **i)** $B_{K,\underline{p}}(.)$ is continuous, positive and non-increasing on $[0, 1]$ **ii)** $B(\theta) = 1\ \forall \theta \leq \underline{p}$, **ii)** There exists $p^* > \underline{p}$ such that a) $B_{K,\underline{p}}$ solves the differential equation (5) on $(\underline{p}, p^*)$, and b) either $p^* = 1$ or $B_{K,\underline{p}}(p^*) = 0$.

That is, instead of imposing a boundary condition on both ends, we leave one free and let p^* be the smallest value of θ for which $B_{K,\underline{p}}(\theta) = 0$ if such a value exists (and $p^* = 1$ otherwise). (See Appendix A)

Step 2: We build on step 1 to construct a continuous function $K \to h(K, \underline{p})$ such that $h(K, \underline{p}) > K$ if $B_{K,\underline{p}}(p^*) > 0$ and $h(K, \underline{p}) < K$ if $p^* < 1$. (See Appendix A and D)

Step 3: We denote by $\beta_{K,\underline{p}} = (\mathcal{C}_2^{K,\underline{p}}, B_{K,\underline{p}})$ the conditional belief associated to the distribution $B_{K,\underline{p}}$, and we define the mapping $\mathbf{w} : (K, \underline{p}) \to v_1(\tau^D, \beta_{K,\underline{p}})$ which associates to any pair $(K, \underline{p})$ the value obtained by player 1 when he defects and when player 2 follows $\beta_{K,\underline{p}}$. (See Appendix C)

Step 4: Assume that player 2 switches to defect at a constant rate K as long as player 1 cooperates, and that player 1 obtains a continuation payoff equal to w whenever he switches to defect and if player 2 has not switched yet. We show that player 1's best response is a threshold strategy $\tau_1^{K,\underline{q}}[\theta]$ for some θ, and we let $\underline{\mathbf{q}}(K, w)$ denote the threshold (See Appendix C).

Step 5: We show that the function $J : (K, \underline{p}) \to (h(K, \underline{p}), \underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p})))$ has a fixed point, applying Brower fixed point theorem to J defined on an appropriate convex and compact set (See Appendix D). ■

[26]A direct proof would consist in showing that there exists a continuous selection $K \to (\underline{p}(K), B(K))$ in the set of invariant and consistent conditional beliefs. This would ensure that $K \to \underline{q}(K)$ is also continuous, and by comparing $\underline{p}(K)$ and $\underline{q}(K)$ for small and large values of K, we would conclude that $\underline{p}(K) = \underline{q}(K)$ for some K. Such a direct proof is possible for the introductory case, where it can be ckecked directly that the distribution $B(K) \equiv B_{T(K)}$ has the desired properties. In the general case however, we do have an explicit expression for the $K-$consistent and invariant distributions, and we could not prove that the continuous selection mentionned above exists.

[27]The functions defined in the Appendices are essentially identical to those presented here, up to a change of variable.

Given the bounds provided in Propositions 5 and 6, it is immediate to check that the transition rate K solution to the conditions **i)** and **ii)** of Proposition 8 satisfies

$$| \ln K | \geq \frac{d}{r^{1/2}} \tag{7}$$

for some constant d independent of K and r. We may now conclude.

Proof of Theorem 1: For any discount rate r, consider $(K, \underline{p}, B)$ satisfying the conditions of Proposition 8. We let $b(.)$ be the density associated with the cumulative distribution B and let $\pi = \int_{\underline{p}}^{1} \theta b(\theta) d\theta$.[28] We also let $\beta_i = (\mathcal{C}_j^{K,\underline{p}}, B)$.

Inequality (7) holds, so we may choose r_0 such that for any $r \leq r_0$ and $K, \underline{p}, B$ defined as above, $K^{1/2} \leq r$.

We set $\Theta_i = [\underline{p}, 1]$. Recall that $\bar{\Theta}_i = \{0,1\} \times \Theta_i$, and that a type contingent strategy $\boldsymbol{\tau}_i$ associates to each type $\bar{\theta}_i = (s_i, \theta_i)$ a strategy $\boldsymbol{\tau}_i[\bar{\theta}_i]$ to be followed in case $s_i = 1$. We consider the type-contingent strategy $\boldsymbol{\tau}_i$ defined by

$$\forall \bar{\theta}_i = (s_i, \theta_i), \qquad \boldsymbol{\tau}_i[\bar{\theta}_i] = \boldsymbol{\tau}_i^{K,\underline{p}}[\theta_i]$$

We choose the distribution $\xi \in \Delta(\bar{\Theta}_1 \times \bar{\Theta}_2)$ as follows: first draw (θ_1, θ_2) from the distribution $\xi^0 \in \Delta(\Theta_1 \times \Theta_2)$ defined by

$$\xi^0(\theta_1, \theta_2) = b(\theta_1) b(\theta_2).$$

Then draw the state profile (s_1, s_2) according to the following distribution:

$s_1 \backslash s_2$	1	0
1	$\theta_1\theta_2$	$(1-\theta_1)\theta_2$
0	$\theta_1(1-\theta_2)$	$(1-\theta_1)(1-\theta_2)$

By construction, any player i of type $\bar{\theta}_i = (1, \theta_i)$ has a belief $\bar{\beta}_i = (\theta_i, \beta_i)$ with $\beta_i = (\mathcal{C}_j^{K,\underline{p}}, B)$ concerning the other player's state and strategy. And from Proposition 8 it is optimal for player i to follow $\boldsymbol{\tau}_i[\bar{\theta}_i]$.

When his type is $(1, \theta_i)$, player i's equilibrium payoff is equal to

$$\theta_i v_1(\tau^{K,\underline{p}}(\theta_i), \beta) + (1-\theta_i) v_1(\tau^{K,\underline{p}}(\theta_i), \tau^D),$$

which, by Proposition 7 and since $K^{1/2} \leq r$, is at least equal to $\theta_i + O(r \mid \ln r \mid)$.

Ex ante, player i turns out to be in physical state $s_i = 0$ with probability $\int_{\underline{p}}^{1} (1 - \theta_2) b(\theta_2) d\theta_2 = 1 - \pi$. Player i's expected equilibrium payoff is thus at least equal to $\pi[\pi + O(r \mid \ln r \mid)]$. Since $1 - \pi \leq O(K^{1/2})$ (from Proposition 5), and since $K^{1/2} \leq r$, we finally get that for r_0 small enough, player i's expected equilibrium payoff is as close as we want from 1. ■

[28] Note that B is differentiable on $(\underline{p}, 1)$ because it is continuous and solves the differential equation (5).

Extension 1 The first extension we wish to point out is that our result remains valid even when we restrict attention to correlated equilibria where both players are in physical state 1 at the start of the game. That is, each player may start by cooperating if he wants to.

We choose r, K, $\underline{p}$, B as in Theorem 1, and now choose r_0 small enough so that $K^{1/4} \mid \ln K \mid \leq r^3$. We define $\bar{\pi} = \int_{\underline{p}}^1 b(\theta)\frac{1-\theta}{\theta}$. It follows from Proposition 5 that $\bar{\pi} \leq dK^{1/2} \mid \ln K \mid$ for some constant d. We set $\Theta_i = \{0\} \cup [\underline{p}, 1]$. We consider the type-contingent strategy $\boldsymbol{\tau}_i$ defined by

$$\boldsymbol{\tau}_i[0] = \tau^D \qquad \text{and} \qquad \forall \theta_i \geq \underline{p}, \qquad \boldsymbol{\tau}_i[\theta_i] = \boldsymbol{\tau}_i^{K,\underline{p}}[\theta_i].$$

And we define a distribution $\xi \in \Delta(\Theta_1 \times \Theta_2)$ as follows:

$$\begin{aligned} \xi(\theta_1, \theta_2) &= ab(\theta_1)b(\theta_2) \text{ for all } \theta_1 \geq \underline{p} \text{ and } \theta_2 \geq \underline{p}, \\ \xi(0, \theta_2) &= ab(\theta_2)\frac{1-\theta_2}{\theta_2} \text{ for all } \theta_2 \geq \underline{p} \\ \xi(0,0) &= adK^{1/4} \end{aligned}$$

where a solves $a(1 + 2\bar{\pi} + dK^{1/4}) = 1$. Note that this implies $a \simeq 1 - O(K^{1/4})$.

By construction, any player i of type $\theta_i \geq \underline{p}$ believes that player j starts by defecting with probability θ_i and follows $\beta_i = (\mathcal{C}_j^{K,\underline{p}}, B)$ otherwise. It is therefore optimal for player i to follow $\boldsymbol{\tau}_i^{K,\underline{p}}[\theta_i]$. When player i's type is 0, he believes that his opponent is also of type 0 with probability at least equal to $1 - K^{1/4} \mid \ln K \mid$.

Given our choice of r_0, $K^{1/4} \mid \ln K \mid \leq r^3$, and by Lemma 4 (see Appendix E), it is then optimal for player 1 to switch to defect from the start.

Finally, as in the proof of theorem 1, when his type is $\theta_i \geq \underline{p}$, player i's equilibrium payoff is at least equal to $\theta_i + O(r \mid \ln r \mid)$. Ex ante, player i's expected equilibrium payoff is thus at least equal to $a \int_{\underline{p}}^1 \theta b(\theta) d\theta + O(r \mid \ln r \mid)$, which is as close as we want from 1 for r_0 small enough.

Extension 2. Throughout the paper, we have assumed that defections are irreversible. We wish to show here that cooperation may be sustained even if players may revert to cooperation after some (long enough) time. We denote by $G(T)$ the game where players may revert to cooperation after defecting for a length of time at least equal to T, and we look for correlated equilibria where players never revert to cooperation, even though they have the option to do so. We assume that when a player defects, he receives no signals (whether his opponent defects or not).

Proposition 9 *For any given* μ, λ, b, c *with* $\mu > \lambda \geq 0$ *and* $c > 1 > 0 > b$*, and for any* $\varepsilon > 0$*, there exists* r_0*,* $r_1 < r_0$ *and* T_0 *such that for any* r*,* $0 < r_1 \leq r \leq r_0$*, and for any* $T \geq T_0$*, the game* $G(T)$ *has a correlated equilibrium* $(\Theta, \xi, \boldsymbol{\tau})$ *such that* $v_i(\boldsymbol{\tau}, \xi) \geq 1 - \varepsilon$ *for each player* $i = 1, 2$*.*

The intuition is very simple. After having defected during T_0 units of time, player 1 knows that player 2 will have almost surely switched to defect. At this point, trying to learn if indeed player 2 has switched to defect or not would be too costly. The proof actually shows that we may choose T_0 small compared to $1/r$ when c gets large. The details are in Appendix E.

6 Discussion

Our main result shows that for any given sufficiently small discount rate, one can find a correlation device that allows players to sustain cooperation. One weakness of our result is thus that we do not provide any answer concerning the case where no such correlation device is available or where players start out with a different correlation device. We wish to make some comments concerning these two cases and their relationship to the case we have analyzed.

As mentioned in Section 2, for any initial belief $\bar{\beta}_1^0 = (p_1^0, \beta_1^0)$ of player 1 concerning player 2's physical state and strategy, in any continuation game, player 1's continuation beliefs are the form $\bar{\beta}_1^t = (p_1^t, \beta_1^t)$ where $p_1^t = p_1^t(\omega_1^t, \bar{\beta}_1^0)$ and $\beta_1^t = \beta_1^t(\beta_1^0)$; that is, continuation conditional beliefs evolve deterministically. Besides, best responses would remain of the form "Continue to cooperate as long as $p_1^t(\omega_1^t, \bar{\beta}_1^0) > \underline{p}^t$," the only difference with the case analyzed in the paper being that the threshold $\underline{p}^t$ may vary over time.

To illustrate the consequence of these observations, consider the simple introductory case. The strategies used by players in equilibrium would be of the form:

> *At any date* t*, switch to defect if the lapse of time* θ_i^t *since the last signal received exceeds* $T(t)$*.*

Assume that player 2 follows such a strategy, and denote it by τ_2. The evolution of player 1's conditional beliefs can be derived as follows. Let $\bar{B}(t, \theta) = \Pr_{\tau^C, \tau_2}\{s_2^t = 1, \theta_2^t \leq \theta\}$ be the function that characterizes the (expected) unconditional beliefs, and let $B(t, \theta) = \frac{\bar{B}(t,\theta)}{\bar{B}(t,T(t))}$ be the function that characterizes conditional beliefs. It is easy to check that $\bar{B}(t, \theta)$ solves the following partial differential equation:[29]

$$\frac{\partial \bar{B}}{\partial t} = \mu[\bar{B}(t, T(t)) - \bar{B}(t, \theta)] - \frac{\partial \bar{B}}{\partial \theta} \tag{8}$$

In the paper, we have analyzed the case where $T(t) = T$ for all t. In that case, and for a well chosen initial distribution $B(0, \theta) = B_T(\theta)$, (8) admits a separable solution $\bar{B}(t, \theta) = H(t)B_T(\theta)$.[30]

In contrast, when the function $T(t)$ is not constant, or when one starts with conditional beliefs different from $B_T(\theta)$, the function $B(t, \theta)$ cannot be time invariant. Nevertheless, the analysis of the partial differential equation (8) may prove useful for several reasons:

a) Assume that there exists an equilibrium characterized by a function $T(t)$ which eventually converges (smoothly) to some T. It is easy to show that $B(t, \theta)$ eventually converges to $B_T(\theta)$.[31] Equilibrium behavior and beliefs would thus eventually coincide with those analyzed in the current paper.

b) By analyzing the speed of convergence of $B(t, \theta)$ towards $B_T(\theta)$ when $T(t)$ tends to T, one may be able to show (if the speed of convergence is large enough) that best responses would also remain very close to T. We would thus be address local stability issues concerning the equilibria we have found.

These issues are left for future research.

On the interpretation of our main result Although stated in positive terms, a more cautious interpretation of our result could be given, namely, that sustaining cooperation without public observations requires much more sophistication on the part of the players than sustaining cooperation with public observations.

Indeed, consider the case where monitoring is public. Assume for example that signals are public and arise according to a Poisson process with arrival rate μ when both players cooperate, and with arrival rate $\lambda < \mu$ when one players defects. Let $T = d \mid \ln r \mid$. A possible strategy would consist in reviewing at each date kT, $k \geq 1$ the number of signals received since date $(k-1)T$, and switch to defect if the average number of signals received per unit of time is smaller than $\frac{\mu+\lambda}{2}$. As T increases, the probability of mistakes decreases exponentially (see Lemma 2 in the Appendix). Hence one may choose d large

[29]The term $\mu[B(t, T(t)) - B(t, \theta)]$ corresponds to the flow of entry into $[0, \theta]$ and the term $\frac{\partial B}{\partial \theta}$ corresponds to the flow of exit.

[30]This solution is such that $\frac{H'(t)}{H(t)} = -K$ for some K, $\mu(1 - B(\theta)) - B'(\theta) = -KB(\theta)$, and $B(T) = 1$.

[31]Convergence is easy to show because around T, $B(t, \theta) - B(\theta)$ is close to 0 for t large by construction, and one may then work backward from high values of θ to lower values of θ.

enough (independently of r) so that the probability to trigger the defection when players are actually cooperating is very small (compared to r^2), and so that the probability that the punishment will be triggered when one player defects is very close to 1 (larger than $1 - r^2$). For such a constant d, the strategy profile considered is in equilibrium.

Admittedly, some cooperation can be sustained when monitoring is private, but at what cost? First, an appropriate correlation device has to be implemented. And even when this is possible, the strategies that players ought to follow in equilibrium are quite demanding, as they require to compute the path $q^{\omega_2,\theta_2}(t)$. Finally, if one wants to forget about implementing a correlation device, then an equilibrium supporting cooperation (if it exists) would impose even greater sophistication on the strategies used.

Our paper may therefore also contribute to understanding why sustaining cooperation with imperfect private observations is difficult.

On the relationship with the reputation literature. As mentioned earlier, dynamic games with private and imperfect monitoring are inherently games of incomplete information. Here the incomplete information bears on whether one's opponent has already switched to defect. Interpreting a player who has already defected as a *bad type*, one may compare our work with the reputation literature.

In our game, each player (endogenously) becomes a bad type with a small but positive probability. Each player thus realizes that even when he cooperates, there is a positive probability that he will be mistaken for a bad type. But he also realizes that if he starts defecting, he will very rapidly acquire a reputation for being a bad type. And in equilibrium, this provides him with incentives to cooperate. Players are thus deterred from defecting in equilibrium, because *they want to avoid being mistaken for a bad type* (For a similar insight in a different context -players exogenously become bad types with positive probability- see Mailath and Samuelson (1998)).

7 Concluding comments

In games with imperfect public monitoring, public signals allow players to coordinate their strategies in any continuation game.[32] In the absence of public signals or communication (as in Compte 1998 or Kandori and Matsushima 1998), coordinating play in any continuation game is not possible.

[32] The literature on repeated games with imperfect public monitoring includes Green and Porter (1984), Abreu Pearce and Stacchetti (1990), Fudenberg Levine and Maskin (1994) and Fudenberg and Levine (1994).

Coordination however may not be required to support cooperation. What is needed is that deviations be detected and punished. In situations where a player's actions have an effect on the distribution over his opponent's signals, deviations can be detected and we just have to provide each player with incentives to use their observations appropriately so that deviators be punished.

The main difficulty is that compared to the case of public monitoring, we have much less freedom in choosing how players condition play on past observations, because a player will use past observations *if and only if* these observations convey information about his opponent's future behavior. This has several implications:

1) In a setting where observations are purely private (i.e. conditionally independent), a player's observations may only convey information about his opponent's future behavior if the two following conditions hold:

1. - there is (in equilibrium) some uncertainty about the actions played by the opponent;
2. - there is (in equilibrium) some correlation between the opponent's past play and the opponent's future behavior.

In papers where players are constrained to play pure strategies, uncertainty about other players' actions is absent, and negative result are found. (see Matsushima (1990) and Bagwell (1995)). In papers where mixed strategies are allowed or where extra uncertainty is introduced, positive results may be found. (see Sekiguchi (1997), Bhaskar and van Damme (1997), Mailath and Morris (1997) and van Damme and Hurkens (1997)).

In our paper, the probability that a player switches to defect is small in equilibrium, but it is important that it never vanishes. If it did, players would soon pay little attention to their signals, and incentives to cooperate would not be preserved.

2) The signals received by a player, say player 1, will convey information about the actions played by player 2. However this information will be interpreted in the light of the distribution over the set of paths of actions that may be played by 2 in equilibrium. If player 2 is supposed to follow a trigger strategy (and if signals are conditionally independent), then player 1 will only try to distinguish between two events: Has player 2 already switched to defect, or Is he still cooperating? This constrains the set of tests that players would be willing to carry out in equilibirum, and this constraint was shown to the failure of cooperation once short-run defections are permitted (see Compte (1996)).

3) The third implication is that as soon as there is some uncertainty about past play and some correlation between past and future play, *all* the signals received by a player will convey some information about future play, hence will be used in equilibrium. Equilibrium strategies are thus functions that are likely to be increasingly complex over time, and finding equilibria is in general a very difficult task. One contribution of this paper has been to highlight this difficulty in a very simple setting, and to show how, by looking at small perturbations of the game, one could make the problem tractable again. Whether this technique can be extended to more general settings, in particular to ones where incentives have to be checked for various histories of past actions (and not only for the histories that consist of cooperation at all past dates) is left for future research.

Appendix A

The purpose of this Appendix is to show the existence of an invariant and consistent conditional belief (Proposition 3). We concentrate on the analysis of the differential equation (5). Consider the change of variable

$$\theta \to y_K(\theta) = \frac{\mu - K}{\mu - \lambda + K} \ln \frac{(\mu - \lambda)(\bar{p} - \theta)}{K\theta}.$$

Note that $y_K(1) = 0$, and that, given the definition of $\phi_{n,K}$, we have:

$$\phi_{n,K}(1, \frac{y_K(\theta)}{\mu - K}) \equiv \theta.$$

The ratio $\frac{y_K(\theta)}{\mu - K}$ thus corresponds to the time necessary for a player's type (s, θ) to go from $(1,1)$ to $(1, \theta)$. We also let $\psi_K(y)$ satisfy:[33]

$$\psi_K(y_K(\theta)) = y_K(\phi_s^{-1}(\theta)).$$

It is easy to check that equation (5) and boundary conditions (6) become respectively:

$$B(y) + B'(y) = \frac{\mu}{\mu - K} B(\psi_K(y)) \tag{9}$$

[33]The function $\psi_K(y)$ is explicitly given by:

$$\psi_K(y) = y + \frac{\mu - K}{\mu - \lambda + K}[\ln \frac{\mu}{\lambda} + \ln(1 - \frac{\mu - \lambda}{\mu} e^{-\frac{\mu - \lambda + K}{\mu - K} y})].$$

Note that the time $T(\theta) = \frac{\psi_K(y_K(\theta)) - y_K(\theta)}{\mu - K}$ corresponds to the lapse of time necessary to compensate for the effect of a signal when a player's initial type is $(1, \phi_s^{-1}(\theta))$:

$$\phi_{n,K}(\theta, T(\theta)) \equiv \phi_s^{-1}(\theta)$$

and

$$B(0) = 0 \text{ and } B(y) = 1 \text{ for all } y \geq \bar{y} \tag{10}$$

where $\bar{y} = y_K(\underline{p})$.

For any $\bar{y} > 0$ and $K \in (0, \mu)$, one can construct a solution $S(y, \bar{y}, K)$ to the differential equation (9) iteratively by considering the successive intervals $[\bar{y}^{(1)}, \bar{y}], \ldots, [\bar{y}^{(n)}, \bar{y}^{(n-1)}], \ldots$.where $\psi_K^{(n)}(\bar{y}^{(n)}) = \bar{y}$.[34,35] The main difficulty is to show that we can find $\bar{y}$ for which $\lim_{y \searrow 0} S(y, \bar{y}, K)$ is well defined and equal to 0. In what follows, we let

$$\mathbf{y}^*(\bar{y}, K) = \inf\{y \mid y \geq 0 \text{ and } S(y', \bar{y}, K) > 0 \ \forall y' \geq y\}$$

We define the map:

$$\widehat{S} : (y, \bar{y}, K) \rightarrow \begin{cases} S(y, \bar{y}, K) \text{ for } y > \mathbf{y}^*(\bar{y}, K) \\ \lim_{y \searrow y^*} S(y, \bar{y}, K) \text{ for } y \leq \mathbf{y}^*(\bar{y}, K) \end{cases} \tag{12}$$

We have the following Lemma:

Lemma 1 i) *For any* $(\bar{y}, K) \in \Re^+ \times (0, \mu)$*,* $\widehat{S}(., \bar{y}, K)$ *is continuous, positive and non-decreasing on* $\Re^+$*, and the mapping* $(\bar{y}, K) \rightarrow \widehat{S}(., \bar{y}, K)$ *is continuous (in the norm sup topology):*
ii) *The mappings* $\mathbf{y}^* : (\bar{y}, K) \rightarrow \mathbf{y}^*(\bar{y}, K)$ *and* $\mathbf{s} : (\bar{y}, K) \rightarrow \widehat{S}(\mathbf{y}^*(\bar{y}, K), \bar{y}, K)$ *are well defined and continuous on* $\Re^+ \times (0, \mu)$*.*
iii) *Choose* $a < 1$*. There exists* K_0 *such that*

$$\forall K \in (0, K_0), \ \forall y \geq \bar{y} - a \mid \ln K \mid, \widehat{S}(y, \bar{y}, K) > 1 - \frac{K^{1-a}}{\mu}.$$

iv) *There exists* K_0 *such that:*

$$\forall K \in (0, K_0), \ \exists \bar{y}_0, \ \forall \bar{y} > \bar{y}_0, \ \mathbf{y}^*(\bar{y}, K) > 0.$$

Proof. In what follows, we fix $(\bar{y}, K)$ and let $y^* = \mathbf{y}^*(\bar{y}, K)$. For any $\nu > 0$, the set $\mathcal{N}_\nu(\bar{y}, K)$ denotes a ball of size ν around $(\bar{y}, K)$.

---

[34] The superscript (n) denotes the n^{th} iterate, that is, for any function f, $f^{(1)} = f$ and $f^{(n)} = f \circ f^{(n-1)}$.

[35] At each step, multiply each side of equation (9) by e^y and integrate on both sides. We obtain, for all $y \in [\bar{y}^{(n+1)}, \bar{y}^{(n)}]$:

$$S(y, \bar{y}, K) = e^{\bar{y}^{(n)} - y} S(\bar{y}^{(n)}, \bar{y}, K) - \int_y^{\bar{y}^{(n)}} e^{z-y} S(\psi_K(z), \bar{y}, K) dz \tag{11}$$

i) By construction, $S(.,\bar{y},K)$ is continuous on $(0,+\infty)$ and differentiable on $(0,\bar{y})$. Besides, it is easy to check that for any $y \in (y^*,\bar{y})$, we have[36]

$$0 < \frac{\partial}{\partial y} S(y,\bar{y},K) \leq \frac{\mu}{\mu - K}. \tag{13}$$

So the limit $\lim_{y \searrow y^*} S(y,\bar{y},K)$ is well defined (even when $y^* = 0$), and $\widehat{S}(.,\bar{y},K)$ is continuous, positive and non-decreasing.

For any $y > 0$, $(\bar{y},K) \to \widehat{S}(y,\bar{y},K)$ is continuous and, thanks to inequalities (13) the function $y \to \widehat{S}(y,\bar{y},K)$ has bounded right and left derivatives. It is then standard to prove that $(\bar{y},K) \to \widehat{S}(.,\bar{y},K)$ is continuous in the norm sup topology.[37]

ii) Choose any $\varepsilon > 0$. We let $\bar{s} = \widehat{S}(y^* + \varepsilon, \bar{y}, K)$. By definition of $\mathbf{y}^*$, and thanks to (13), $\widehat{S}(y,\bar{y},K) \geq \bar{s} > 0$ for all $y \geq y^* + \varepsilon$. It follows from **i)** that for any $(\bar{y}',K') \in \mathcal{N}_\nu(\bar{y};K)$ with ν small enough, $\widehat{S}(y,\bar{y},K) \geq \frac{\bar{s}}{2} > 0$ for all $y \geq y^* + \varepsilon$. And thus $\mathbf{y}^*(\bar{y}',K') \leq y^* + \varepsilon$.

If $y^* = 0$, then $\mathbf{y}^*(\bar{y}',K') \geq y^*$ since $\mathbf{y}^*(\bar{y}',K')$ is non-negative by definition of $\mathbf{y}^*$.

If $y^* > 0$, then Equation (9) implies $\frac{\partial S(.,\bar{y},K)}{\partial y} \mid_{y=y^*} > 0$. Hence there exists $\varepsilon' < \varepsilon$ and $\underline{s} < 0$ such that $S(y^* - \varepsilon', \bar{y}, K) = \underline{s}$ and $y^* - \varepsilon' > 0$. Since $(\bar{y},K) \to S(y,\bar{y},K)$ is continuous for any $y > 0$, there exists ν' such that for all $(\bar{y}',K') \in \mathcal{N}_\nu(\bar{y};K)$, $S(y^* - \varepsilon', \bar{y}, K) \leq \frac{\underline{s}}{2} < 0$, in which case $\mathbf{y}^*(\bar{y}',K') \geq y^* - \varepsilon'$.

Continuity of $\mathbf{s}$ follows immediately from that of $\mathbf{y}^*$ and $(\bar{y},K) \to \widehat{S}(.,\bar{y},K)$.

iii) Since $S(y,\bar{y},K) \leq 1$, Equation (9) implies that for all $y \geq y_0$, the function $S(.,\bar{y},K)$ satisfies

$$B'(y) + B(y) \leq \frac{\mu}{\mu - K}.$$

Multiplying by e^y on both sides and integrating, one gets:

$$S(y,\bar{y},K) \geq \frac{\mu}{\mu - K} - \frac{K}{\mu - K} e^{(\bar{y} - y)},$$

[36] Assume by contradiction that there exists $y_0 \in (y^*,\bar{y})$ such that $\frac{\partial S(y,\bar{y},K)}{\partial y} \mid_{y=y_0} = 0$ and consider the largest such y_0. We must have $y_0 < \bar{y}$ because equation (11) implies $\frac{\partial S(y,\bar{y},K)}{\partial y} > 0$ on $[\bar{y}^{(1)},\bar{y})$. Thus for $y \in (y_0,\bar{y})$, $\frac{\partial S(y,\bar{y},K)}{\partial y} > 0$, hence $S(\psi(y_0),\bar{y},K) > S(y_0,\bar{y},K)$. Since $S(.,\bar{y},K)$ satisfies Equation (9), we obtain $\frac{\partial S(y,\bar{y},K)}{\partial y} \mid_{y=y_0} > 0$, which is a contradiction.

The upper bound follows from observing that $S(\psi(y),\bar{y},K) \leq 1$ and $S(y,\bar{y},K) \geq 0$ for all $y \in [y^*,\bar{y}]$.

[37] Let M be the bound on the derivatives. Consider any given $\varepsilon > 0$, and choose a grid $(y_1, ..., y_n)$ of size $\frac{\varepsilon}{2M}$. Since $(\bar{y},K) \to \widehat{S}(y_k,\bar{y},K)$ is continuous for all $k \in \{1,n\}$ and since n is finite, there exists v such that $\mid \widehat{S}(y_k,\bar{y},K) - \widehat{S}(y_k,\bar{y}',K') \mid \leq \varepsilon/2$ for all $k \in \{1,n\}$ and $(\bar{y}',K') \in \mathcal{N}_\nu(\bar{y};K)$. Finally observe that for all y, there exists k such that $\mid y - y_k \mid \leq \frac{\varepsilon}{2M}$ and

$$\mid \widehat{S}(y,\bar{y},K) - \widehat{S}(y,\bar{y}',K') \mid \leq \mid \widehat{S}(y,\bar{y},K) - \widehat{S}(y_k,\bar{y},K) \mid + \mid \widehat{S}(y_k,\bar{y},K) - \widehat{S}(y_k,\bar{y}',K') \mid.$$

The first term on the right hand side is smaller than $M\frac{\varepsilon}{2M}$, and the second term is bounded by $\varepsilon/2$. The sum of these two terms is therefore bounded by ε independently of y and $(\bar{y}',K') \in \mathcal{N}_\nu(\bar{y};K)$.

which implies the desired conclusion.

iv) Observe that $S(.,\bar{y},K)$ satisfies $B(\psi_K(y)) \geq B(y)$ on $[y^*,\bar{y}]$, hence using (9), it also satisfies $B'(y) \geq \frac{K}{\mu-K}B(y)$ for $y \in (y^*,\bar{y})$, which implies that for any $y \in [y^*,\bar{y}]$,

$$S(y,\bar{y},K) \leq e^{-\frac{K}{\mu-K}(\bar{y}-y)} \tag{14}$$

Also, it is easy to check that there exists $K_0 > 0$, $a_0 > 0$ and $A > 0$ such that for all $K \leq K_0$ and $y \geq A \mid \ln K \mid$,[38]

$$\psi_K(y) - y - 1 > a_0 \text{ and } \psi_K'(y) \leq \frac{\mu}{\mu - K} \tag{15}$$

Fix $K \leq K_0$ and $y_0 \geq A \mid \ln K \mid$. We choose $\bar{y} > y_0$ and assume (by contradiction) that $\mathbf{y}^*(\bar{y},K)$ is below y_0 Equation (9) implies that for all $y \geq y_0$, the function $S(.,\bar{y},K)$ satisfies

$$B'(y) \geq \psi_K'(y)B(\psi_K(y)) - B(y).$$

Integrating on both sides, we obtain:

$$1 - B(y_0) \geq \int_{\bar{y}}^{\psi_K(\bar{y})} B(y)dy - \int_{y_0}^{\psi(y_0)} B(y)dy,$$

which further implies (since $B(y) = 1$ for all $y \geq \bar{y}$),

$$B(y_0) \leq -[\psi(\bar{y}) - \bar{y} - 1] + \int_{y_0}^{\psi(y_0)} B(y)dy - B(y_0) \tag{16}$$

When $\bar{y}$ gets large, the term in brackets remains larger than a_0 while the integral vanishes (because $B(y)$ vanishes thanks to inequality (14) and because $\psi(y_0) - y_0$ remains bounded above). Thus $B(y_0) < 0$, contradicting the assumption $\mathbf{y}^*(\bar{y},K) \leq y_0$ So we must have $\mathbf{y}^*(\bar{y},K) > y_0$ when $\bar{y}$ is large enough. ■

Lemma 1 has the following Corollary.

Corollary 1 *There exists $K_0 > 0$ such that for any $K \in (0,K_0]$, there exists $\bar{y}$ and $S(.,\bar{y},K)$ solving (9) with boundary conditions (10).*

Proof. For any given $K \in (0,K_0]$, consider the mapping $h : \bar{y} \to \mathbf{y}^*(\bar{y},K) - \mathbf{s}(\bar{y},K)$. It is continuous, negative when $\bar{y} < \frac{1}{2} \mid \ln K \mid$ and positive for $\bar{y}$ large enough. So there must exist $\bar{y}$ for which $h(\bar{y}) = 0$. Since $\mathbf{y}^*(\bar{y},K)$ and $\mathbf{s}(\bar{y},K)$ cannot be simultaneously different from 0, we have must $\mathbf{y}^*(\bar{y},K) = \mathbf{s}(\bar{y},K) = 0$. ■

Corollary 1 implies that for any $K \in (0,K_0)$, a solution to the differential equation (5) with boundary conditions (6) exists. Proposition 4 then implies that for any such K, an invariant conditional belief consistent with K exists.

[38] Using the expression of ψ_K given footnote 33, note that when K tends to 0, $\psi_K(y) - y$ tends to $\frac{\mu}{\mu-\lambda}\ln\frac{\mu}{\lambda} = \frac{|\ln\lambda/\mu|}{1-\lambda/\mu}$, which is strictly larger than 1 because $\lambda/\mu < 1$.

Appendix B

The purpose of this Appendix is to give some properties satisfied by the solutions of (5-6). This Section and the next one will make use of the following standard result in statistics (see Compte 1996 for further details).

Lemma 2 *Let $h(x) = x \log x - (x - 1)$. For a Poisson process with arrival rate γ, we have:*

$$\begin{aligned} \Pr(\frac{n_t}{\gamma t} &\leq x) \leq \exp -\gamma t h(x) \text{ if } 0 < x < 1, \text{ and} \\ \Pr(\frac{n_t}{\gamma t} &\geq x) \leq \exp -\gamma t h(x) \text{ if } x > 1. \end{aligned}$$

We now prove Proposition 5.

Proof of Proposition 5.

Throughout this proof, we assume that player 1 cooperates and that player 2 follows $\beta = (\mathcal{C}_2^{K,\underline{p}}, B)$. Recall that the latter assumption means that if player 2's initial type is $(1, \theta_2^0)$, then player 2 is still in physical state $s_2 = 1$ at date t_0 if and only if $q^{\omega_2, \theta_2^0}(t) > \underline{p}$ for all $t < t_0$.

We start with some notation. We let $\widetilde{p} = 1 - K \mid LogK \mid$. When player 2 receives no signals, her type is driven down. We let $\widetilde{t}$ denote the time it would take for player 2's type to reach $(1, \underline{p})$ starting from $(1, \widetilde{p})$. Formally, $\widetilde{t}$ solves[39]

$$\phi_{n,K}(\widetilde{p}, \widetilde{t}) \equiv p.$$

At each signal received, player 2's type jumps upward, so signals lengthen the time it takes for player 2's type to reach $(1, \underline{p})$. Formally, we define[40]

$$\widetilde{T} \equiv \max_t \{\phi_{n,K}[\phi_s(p), t] \geq p, \forall p < \phi_s^{-1}(\widetilde{p})\}.$$

That is, if player 2's type at t^- is $(1, p)$ and if he receives a signal at t that keeps him below $(1, \widetilde{p})$, then it takes a lapse of time at least equal to $\widetilde{T}$ for player 2's type to get back to its previous level. The time $\widetilde{T}$ may thus be interpreted as the lapse of time necessary to cancel out the effect of good signal (when $\phi_s(p) \leq \widetilde{p}$). It is easy to check that

$$\widetilde{T} = \frac{\ln \frac{\mu}{\lambda}}{\mu - \lambda} + O(\mid LogK \mid^{-1}). \tag{17}$$

Note that $\frac{\mu}{\lambda} > 1$, so $\widetilde{T}$ is larger than $1/\mu$: the time necessary to cancel out the effect of a good signal is larger than the average time between two consecutive signals. Thus on

[39] Using the notation introduced in Appendix A, we also have $\widetilde{t} = \frac{\bar{y} - \widetilde{y}}{\mu - K}$, where $\widetilde{y} = y_K(\widetilde{p})$.

[40] Using the notation introduced in Appendix A, we could also write $\widetilde{T} = \max_{y \geq \widetilde{y}} \frac{\psi_K(y) - y}{\mu - K}$.

average, player 2's type tends to be pushed up above $\widetilde{p}$. (This observation is key to our result.)

Consider the probability

$$\pi = \Pr_{\tau^C,\beta_K}(s_2^{t_0+\widetilde{T}} = 0, s_2^{t_0} = 1 \mid s_2^0 = 1)$$

that player 2 switches to defect between date $t_0 - \widetilde{T}$ and t_0 when players conform to the profile (τ^C, β). Under β, player 2 switches to defect at constant rate K by assumption, thus we have:

$$\pi = e^{-K(t_0-\widetilde{T})}(1 - e^{-K\widetilde{T}})$$

Now consider all the paths $q^{\omega_2,\theta_2^0}(t)$ that exit from $[\underline{p}, 1]$ during $[t_0 - \widetilde{T}, t_0]$. Among these paths, we distinguish between those for which $q^{\omega_2,\theta_2^0}(t)$ remains below $\widetilde{p}$, and those for which $q^{\omega_2,\theta_2^0}(t)$ went above $\widetilde{p}$ at least once. We denote by π_1 and π_2 the resulting probabilities $(\pi = \pi_1 + \pi_2)$. We are now going to derive bounds on π_1 and π_2.

Let n^t denote the number of signals received by player 2 up to date t. In order to exit from $[\underline{p}, 1]$ during $[t_0 - \widetilde{T}, t_0]$ without going above $\widetilde{p}$, the number of signals received by player 2 should satisfy $n^{t_0}\widetilde{T} \le t_0$, implying that

$$\pi_1 \le \Pr_{\tau^C,\tau^C}(n^{t_0} \le \frac{t_0}{\widetilde{T}}) \tag{18}$$

Similarly, in the event where $t_0 - s$ is the last date before t_0 where $q^{\omega_2,\theta_2^0}(t) = \widetilde{p}$, player 2's type will reach $\underline{p}$ during $[t_0 - \widetilde{T}, t_0]$ only if $\widetilde{T}n^s + \widetilde{t} \le s$, implying that:

$$\pi_2 \le \sup_s \{\Pr_{\tau^C,\tau^C}(n^s \le \frac{s - \widetilde{t}}{\widetilde{T}})\} \tag{19}$$

Let $\beta = \frac{1}{\mu\widetilde{T}}$ and $z = \frac{s-\widetilde{t}}{\mu s\widetilde{T}}$. Applying Lemma 2 yields

$$\pi \le e^{-\mu t_0 h[\beta]} + e^{-\mu\beta\widetilde{t}\min_z \frac{h(z)}{\beta - z}} \tag{20}$$

Choosing $t_0 = \frac{2}{\mu h[\beta]} \mid \ln K \mid$ ensures that the first term is comparable to K^2. Besides let $T^* = \frac{\ln\frac{\mu}{\lambda}}{\mu-\lambda}$ and $\beta^* = \frac{1}{\mu T^*}$. It is easy to check that $\min_z \frac{h(z)}{\beta^*-z} = \ln\frac{\mu}{\lambda}$.[41] It follows that $\min_z \frac{h(z)}{\beta^*-z} = \ln\frac{\mu}{\lambda} + O(\mid LogK \mid^{-1})$, and inequality (20) implies

$$K \le \gamma e^{-(\mu-\lambda)\widetilde{t}}. \tag{21}$$

for some constant γ independent of K. Checking that $\underline{p} \ge \frac{d_1}{|\ln K|}$ for some constant d_1 independent from K is then immediate.[42]

[41] $\frac{h(z)}{\beta^*-z}$ is minimum for z^* satisfying $h'(z^*)(\beta^* - z^*) + h(z^*) = 0$, which is equivalent to $\beta^* = \frac{z^*-1}{\ln z^*}$, or $z^* = \frac{\lambda}{\mu}$. Thus we obtain $\frac{h(z^*)}{\beta^*-z^*} = -h'(z^*) = -\ln z^* = \ln\frac{\mu}{\lambda}$, as desired.

[42] Indeed, inequality (21) implies $K\widetilde{t} \le d_0$ for some constant d_0, and by definition of $\widetilde{t}$, we have $\frac{\bar{p}}{\underline{p}} - 1 \equiv (\frac{\bar{p}}{\widetilde{p}} - 1)e^{(\mu-\lambda)\widetilde{t}+K\widetilde{t}}$.

In order to prove the second part, we choose $p_1 = 1 - K^{1/2}$, and we show that $B(p_1) \geq 1 - K^{1/2}$. The argument is very similar to the one just developed.. Consider the probability

$$\Pi \equiv \Pr_{\tau^C,\beta}(s_2^{t_0} = 1, q^{\omega_2,\theta_2^0}(t_0) \in (\underline{p}_K, p_1) \mid s_2^0 = 1)$$

that player 2's type at date t_0 is $(1, \theta_2^{t_0})$ with $\theta_2^{t_0} \in (\underline{p}, p)$. By definition of β, we must have:

$$\Pi = e^{-Kt_0}(1 - B(p_1))$$

Any path $q^{\omega_2,\theta_2^0}(.)$ that ends in the interval $(\underline{p}, p_1)$ at date t_0 either remains below $\widetilde{p}$ at all dates $t < t_0$ or crosses $\widetilde{p}$ at least once. As before, we may decompose the probability Π in two terms, depending on whether the paths $\{q^{\omega_2,\theta_2^0}(.)\}_{\omega_2,\theta_2^0}$ that end in the interval $(\underline{p}, p_1)$ at date t_0 remain below $\widetilde{p}$ at all dates $t < t_0$ or cross $\widetilde{p}$ at least once. We denote the corresponding probabilities by Π_1 and Π_2 respectively.

The first event requires that player 2 receives at most $\frac{\widetilde{t}+t_0}{\widetilde{T}}$ during these t_0 periods, implying that

$$\Pi_1 \leq \Pr_{\tau^C,\tau^C}(n^{t_0} \leq \frac{t_0 + \widetilde{t}}{\widetilde{T}})$$

Let $\widetilde{t}_1$ solve $\phi_{n,K}(\widetilde{p}, \widetilde{t}_1) \equiv p_1$.The second event requires that player 2 receives at most $\frac{s-\widetilde{t}_1}{T^*}$ signals since the last date $t_0 - s$ for which $q^{\omega_2,\theta_2^0}(t_0 - s) = \widetilde{p}$, implying that

$$\Pi_2 \leq \sup_s \{\Pr_{\tau^C,\tau^C}(n_s \leq \frac{s - \widetilde{t}_1}{\widetilde{T}})\}$$

As above, by applying Lemma 2, we obtain that there exists a constant γ such that

$$1 - B(p) \leq \gamma e^{-(\mu-\lambda)\widetilde{t}_1},$$

which, given the definition of $\widetilde{t}_1$ yields the desired result. ■

Appendix C

Consider a pair $(K, \underline{p})$, and let $\bar{y} = y_K(\underline{p})$. The function $\widehat{S}(., \bar{y}, K)$ (see Appendix A) defines a cumulative distribution on $[0, \bar{y}]$.[43] Let $B_{K,\underline{p}}$ denote the induced distribution on $[\underline{p}, 1]$. The pair $\beta_{K,\underline{p}} = (\mathcal{C}_2^{K,\underline{p}}, B_{K,\underline{p}})$ defines a mixed strategy for player 2 Now define:

$$\mathbf{w} : (K, \underline{p}) \to \mathbf{w}(K, \underline{p}) = v_1(\tau^D, \beta_{K,\underline{p}})$$

[43]This cumulative distribution admits a continuous density over $(y^*, \bar{y})$, with possibly a mass point on y^* in case $\widehat{S}(y^*, \bar{y}, K) > 0$ (and $y^* = 0$).

This function associates to each pair $(K, \underline{p})$ the payoff obtained by player 1 when he defects and player 2 follows $\beta_{K,\underline{p}}$.

Finally, consider the hypothetical case where **i)** player 2 switches to defect at a constant rate K as long as player 1 cooperates and **ii)** player 1 obtains a payoff equal to w when he switches to defect and player 2 still cooperates at the time he switches to defect.

We let $V(\eta, K, w)$ denote the payoff obtained by player 1 when he follows the strategy η and the conditions **i**) and **ii**) above hold.[44] And we define the function $\underline{\mathbf{q}}(.,.)$ by:

$$\underline{\mathbf{q}}(K, w) = \min\{1, \inf_{\{\eta, V(\eta,K,w)>w\}} \frac{-v_1(\eta, \tau^D)}{V(\eta, K, w) - v_1(\eta, \tau^D) - w}\}$$

if there exists η such that $V(\eta, K, w) > w$, and $\underline{\mathbf{q}}(K, w) = 1$ otherwise.

Lemma 3 **i)** *The function* $\mathbf{w}$ *and* $\underline{\mathbf{q}}$ *are continuous in both their argument;* **ii)** *There exists* K_0 *and a constant* d_0 *(independent or* r, K *and* $\underline{p}$*) such that for any* $K \leq K_0$*,*
a) $\mathbf{w}(K, \underline{p}) \leq d_0[r y_K(\underline{p}) + \exp -2y_K(\underline{p})]$
b) there exists η^* *such that* $V(\eta^*, K, w) \geq \frac{r}{r+K} - d_0 r \mid \ln r \mid$ *for all* $w \geq 0$*, and* $v_1(\eta^*, \tau^D) \geq -d_0 r \mid \ln r \mid$*.*
c) if $K \leq r$ *and* $w < 1/3$*, then* $\underline{\mathbf{q}}(K, w) \leq d_0 r \mid \ln r \mid$*.*
iii) *Finally, consider a transition rate* $K \in (0, K_0]$ *and a pair* $(\underline{p}, B)$ *that solves (5) and (6). If player 1's initial belief is described by* $(p_0, \beta_{K,\underline{p}})$*, then a best response for player* 1 *is to follow the threshold strategy* $\boldsymbol{\tau}^{K,\underline{q}}(1, p_0)$ *where* $\underline{q} = \underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p}))$*.*

We first check that Proposition 7 is an immediate corollary of Lemma 3.

Proof of Proposition 7:

Choose $K \leq r$ and $(\underline{p}, B)$ solution to (5) and (6).

i) Proposition 5 implies $\underline{p} \geq \frac{d_1}{|\ln K|}$, hence $y_K(\underline{p}) \leq D_1 \mid \ln K \mid$ for some constant D_1. Result **iii)** in Lemma 1 implies $y_K(\underline{p}) > \frac{|\ln K|}{2}$ (otherwise the boundary condition $B(1) = 0$ would not be satisfied). It follows from 3 **ii)**a) that $\mathbf{w}(K, \underline{p}) \leq d_3 r \mid \ln K \mid$ for some constant d_3.

ii) The existence of η^* is immediate.

iii) Let us show that

$$\mathbf{q}(K, \underline{p}) \leq \max(3d_3, d_0) r \mid \ln K \mid . \tag{22}$$

[44] Note that there may exist no strategy for player 2 for which these two conditions hold simultaneously. However, the payoff $V(\eta, K, w)$ obtained by player 1 under these two conditions when he follows η can nevertheless be defined.

If $\max(3d_3, d_0)r \mid \ln K \mid > 1$, then inequality (22) clearly holds. Otherwise, $\mathbf{w}(K, \underline{p}) \leq 1/3$ and $\mathbf{q}(K, \underline{p}) \leq d_0 r \mid \ln r \mid$ by Lemma 3. Hence Inequality (22) holds (because $K \leq r$.) ■

Proof of Lemma 3:

i) Fix $(K, \underline{p})$, and let $w = \mathbf{w}(K, \underline{p})$, $\bar{y} = y_K(\underline{p})$, and $y^* = \mathbf{y}^*(K, \underline{p})$. Let $W(\tau^D, y, K, \underline{p})$ denote the value obtained by player 1 when he defects and when player 2 follows the threshold strategy $\boldsymbol{\tau}^{K,\underline{p}}([y_K]^{-1}(y))$. We have:

$$\mathbf{w}(K, \underline{p}) = \mathbf{s}(K, \underline{p}) W(\tau^D, \mathbf{y}^*(K, \underline{p}), K, \underline{p}) + \int_{\mathbf{y}^*(K,\underline{p})}^{y_K(\underline{p})} W(\tau^D, y, K, \underline{p}) \frac{\partial \widehat{S}(y, y_K(\underline{p}), K)}{\partial y} dy.$$

The functions $W(\tau^D, ., ., .)$, $\mathbf{s}(., .)$ and $\mathbf{y}^*(., .)$ are continuous in all their arguments. So the first term is continuous in $(K, \underline{p})$. Choose $y_0 = y^* + \nu_0$ and $y_1 = \bar{y} - \nu_0$. For ν close enough to 0, the inequality

$$\mathbf{y}^*(K', y_K(\underline{p}')) < y_0 < y_1 < y_K(\underline{p}')$$

holds for all $(K', \underline{p}') \in \mathcal{N}_\nu(K, p)$. Since the function $(\bar{y}, K) \to \widehat{S}(., \bar{y}, K)$ is continuous (see Lemma 1) and since $\widehat{S}$ solves the differential equation (9), the function $(\bar{y}, K) \to \frac{\partial \widehat{S}(., \bar{y}, K)}{\partial y}$ is also continuous with the norm $\mid f \mid = \sup_{y \in [y_0, y_1]} \mid f(y) \mid$. Since W and $\frac{\partial \widehat{S}}{\partial y}$ are bounded, the residual terms (the integral taken over the intervals $[\mathbf{y}^*(K', \underline{p}'), y_0]$ and $[y_1, y_{K'}(\underline{p}')]$) are arbitrarily small when ν and ν_0 are small enough. So the function $\mathbf{w}$ is continuous in $K, \underline{p}$. Continuity of $\underline{\mathbf{q}}(., .)$ is immediate.

ii) Bound on $\mathbf{w}(K, \underline{p})$.

Let $T^* = \frac{\ln \frac{\mu}{\lambda}}{\mu - \lambda}$ and $\bar{y} = y_K(\underline{p})$. Note that $\lambda T^* < 1 < \mu T^*$. Also note that for any y, $\psi_K(y) - y \leq (\mu - K)T^*$. Consider a realization ω_2 and the path $q^{\omega_2, \theta_2^0}(t)$. Recall that player 2 switches to defect when $q^{\omega_2, \theta_2^0}(t) \leq \underline{p}$, or equivalently, $y_K(q^{\omega_2, \theta_2^0}(t)) \geq \bar{y}$.

Let n_t denote the number of signals received by player 2 before date t. Since $\psi_K(y) - y \leq (\mu - K)T^*$ for all y, we have:[45]

$$y_K(q^{\omega_2, \theta_2^0}(t)) - y_K(q^{\omega_2, \theta_2^0}(0)) \geq (\mu - K)t - n_t(\mu - K)T^*.$$

Player 2 still cooperates at t only if $y_K(q^{\omega_2}(t)) \leq \bar{y}$, hence, since $y_K(q^{\omega_2, \theta_2^0}(0)) \geq 0$, only if

$$\frac{n_t}{\lambda t} \geq \frac{1}{\lambda T^*} - \frac{\bar{y}}{\lambda(\mu - K)T^* t}. \tag{23}$$

Let $K_0 = \mu/2$ and assume $K \leq K_0$. Also let $d = \frac{4}{\mu(1 - \lambda T^*)}$ and $\alpha = \frac{1}{2}(\frac{1}{\lambda T^*} + 1) > 1$, and choose $t^* = \max\{d, \frac{2}{\lambda h(\alpha)}\}\bar{y}$ (where $h(.)$ is the function defined in Lemma 2). Inequality (23) implies that player 2 still cooperates at t^* only if $\frac{n_{t^*}}{\lambda t^*} \geq \alpha$. By Lemma 2, this event

[45] Recall that when no signals are received, $y_K(q^{\omega_2, \theta_2^0}(t)) - y_K(q^{\omega_2, \theta_2^0}(0)) = (\mu - K)t$

has a probability at most equal to $e^{-h(\alpha)\lambda t^*} \leq e^{-2\bar{y}}$ Under the complement event, player 1 obtains a payoff at most equal $c(1-e^{-rt^*}) \leq crt^*$. Hence we finally obtain:

$$\mathbf{w}(K,\underline{p}) \leq cr \max\{d, \frac{1}{h(\frac{\alpha}{2})}\} y_K(\underline{p}) + c\exp(-2y_K(\underline{p})).$$

Bound on $\underline{\mathbf{q}}(K,\underline{p})$.

Fix $T>0$. At any date $t^m = mT$, $m \geq 1$, we let N^m denote of signals received by player 1 since date $(m-1)T$. We consider the strategy η^* for player 1 defined as follows: start by cooperating until date t^1. At any date t^m, switch to defect if $\frac{N^m}{T} < \gamma \equiv \frac{\lambda+\mu}{2}$, and continue to cooperate until t^{m+1} otherwise.

By Lemma 2, we have

$$\Pr_{\tau^C,\tau^2=\tau^C}\{\frac{N^m}{\mu T} < \frac{\gamma}{\mu}\} \leq e^{-\mu h(\frac{\gamma}{\mu})T} \text{ and } \Pr_{\tau^C,\tau^2=\tau^D}\{\frac{N^m}{\lambda T} \geq \frac{\gamma}{\lambda}\} \leq e^{-\lambda h(\frac{\gamma}{\lambda})T}.$$

Let $h = \min\{\mu h(\gamma/\mu), \lambda h(\gamma/\lambda)\}$ and choose $T = \frac{2|\ln r|}{h}$. The second inequality implies that if player 2 plays τ^D, then player 1 switches to defect at $t^1 = T$ with probability $1-r^2$, which implies

$$-v(\eta,\tau^D) \leq \mid b \mid rT + r^2. \tag{24}$$

If, for any realization ω_2, player 1 were to switch to defect exactly at the same date $\tau_2(\omega_2)$ as player 2 does, player 1's expected payoff would be equal to $\frac{r}{r+K}$ (since player 2 switches to defect at a constant rate when player 1 cooperates). We show next that player 1's payoff cannot differ from $\frac{r}{r+K}$ by more than $O(r \mid \ln r \mid)$ when he follows η^*.

Let $\bar{T} = \frac{|\ln r|}{r}$ and $\bar{m} = \bar{T}/T$. At each date t^m, either player 2 still cooperates, and player 1 continues to cooperate with probability $1-r^2$, or player 2 has started to defect, and player 1 switch to defect at date t^{k+1} with probability at least equal to $1-r^2$. Thus, a) for any realization $\tau_2(\omega_2) \leq \bar{T}$, player 1 switches to defect between date $\tau_2(\omega_2)$ and $\tau_2(\omega_2)+2T$ with probability at least equal to $(1-r^2)^{\bar{m}} \approx 1-r \mid \ln r \mid$; and b) for any realization $\tau_2(\omega_2) \geq \bar{T}$, player 1 switches to defect after date $\bar{T}$ with probability at least $(1-r^2)^{\bar{m}}$.

Compared to the payoff player 1 would obtain if he could switch to defect exactly at date $\tau_2(\omega_2)$, player 1 looses a payoff at most equal to $2Tr \mid b \mid + O(r \mid \ln r \mid) + O(e^{-r\bar{T}})$. Thus, when $K \leq r$ and $w < 1/3$, the difference $V(\eta, K, w) - w$ is bounded away from 0 (independently of r, K and w). The upper bound on $\underline{\mathbf{q}}$ follows from (24).

iii) Under the assumption that player 1's initial conditional belief is invariant and consistent with K, player 1 obtains a payoff equal to $v_1(\tau^D, \beta_{K,\underline{p}})$ when he switches to defect, under the event $s_2 = 1$. By construction, this payoff is equal to $\mathbf{w}(K,\underline{p})$.

For any strategy η, the payoff $v_1(\eta, \beta_{K,\underline{p}})$ obtained by player 1 thus coincides with $V(\eta, K, w)$. The threshold $\underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p}))$ therefore coincides with the optimal threshold defined in the proof of Proposition 6. ■

Appendix D

Proof of Proposition 8: Our objective is to prove that there exits a pair $(K, \underline{p})$ that solves simultaneously the two following equations

$$\begin{cases} \underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p})) = \underline{p} \\ \mathbf{y}^*(y_K(\underline{p}), K) = \mathbf{s}(y_K(\underline{p}), K) \end{cases}$$

Let $k = \mid \ln K \mid$ and define the functions:

$$\begin{aligned} h(\bar{y}, k) &= k - \mathbf{y}^*(\bar{y}, e^{-k}) + \mathbf{s}(\bar{y}, e^{-k}) \text{ and} \\ z(\bar{y}, k) &= y_K[\underline{\mathbf{q}}(e^{-k}, \mathbf{w}(e^{-k}, \underline{p}))] \text{ where } \underline{p} = (y_{e^{-k}})^{-1}(\bar{y}). \end{aligned}$$

We wish to show that the function $(\bar{y}, k) \to (z(\bar{y}, k), h(\bar{y}, k))$ has a fixed point.

We fix r_0 small and choose any $r \leq r_0$. We set $k_0 = r^{-1/3}$, $k_1 = r^{-2}$, $K_0 = e^{-k_0}$, $y_0 = y_{K_0}(\frac{d_1}{k_0})$, where d_1 is chosen as in Proposition 5, and $y_1 = ak_1$, with $a \in (1/2, 1)$. Choose $A = \frac{2\mu}{\mu - \lambda}$ and consider the compact and convex set F defined by:

$$F = \{(\bar{y}, k), k \in [k_0, k_1],\ \max\{y_0, ak\} \leq \bar{y} \leq \min\{y_1, y_0 + A(k - k_0)\}\}$$

For any $(\bar{y}, k) \notin F$, we consider the projection $\text{Proj}_F(\bar{y}, k)$ on F as described in Figure 2. Finally, we let:

$$J : (\bar{y}, k) \to \begin{cases} (z(\bar{y}, k), h(\bar{y}, k)) & \text{if } (z(\bar{y}, k), h(\bar{y}, k)) \in F \\ \text{Proj}_F(z(\bar{y}, k), h(\bar{y}, k)) & \text{otherwise} \end{cases},$$

The function J is a continuous function from a compact convex set into itself. By Brower fixed point theorem, it admits a fixed point. We will show that there cannot exist a fixed point on the frontier of F.

We consider each segment of the frontier in turn, and for each segment, we consider an element $(\bar{y}, k)$.

1. $\bar{y} = y_1$. (Segment AB in figure 2) Then $k \geq \frac{a}{A} r^{-2}$. Lemma 1 **iii)** implies that $S(\bar{y} - \frac{1}{2}k, \bar{y}, e^{-k}) \geq 1 - \frac{e^{-k/2}}{\mu}$. Thus when player 2 follows $\beta_{K,\underline{p}}$, player 2 continues to cooperate with probability at least equal to $1 - \frac{e^{-k/2}}{\mu}$ during $\frac{1}{2}\frac{k}{\mu}$ periods, even if she does

Figure 1: The dynamics of beliefs: The good news case.

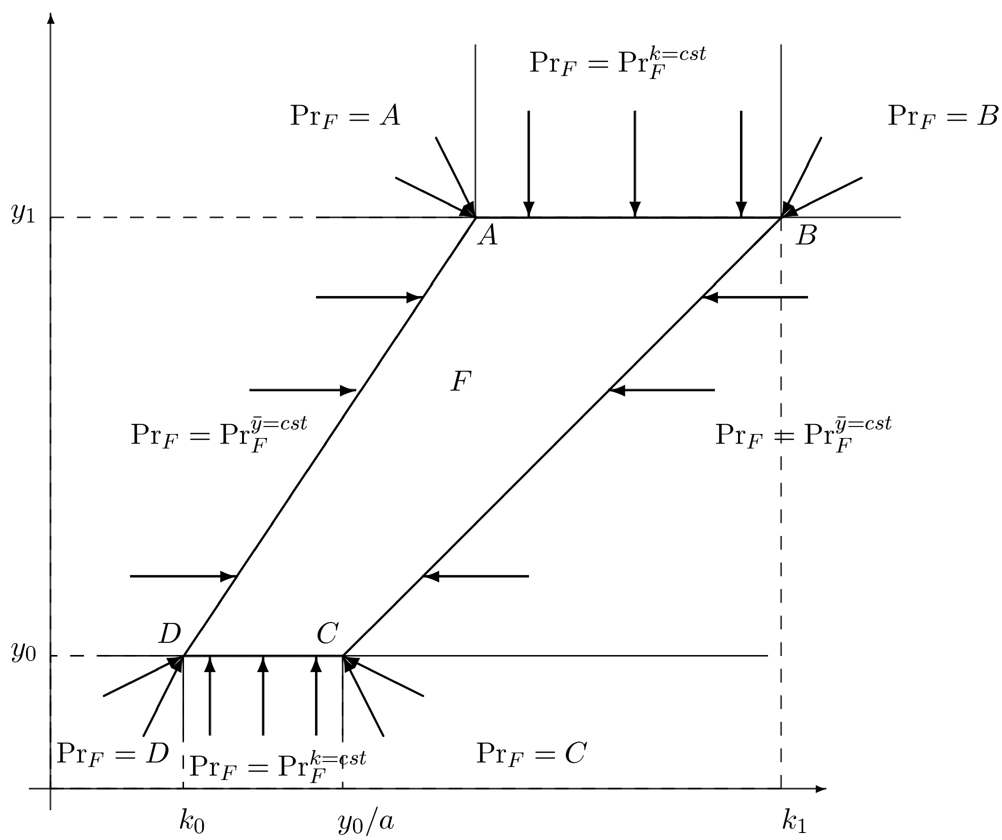

Figure 2: The set F and the projection Pr_F

not receive any signal during that lapse of time.[46] Given that $k \geq \frac{a}{A} r^{-2}$, player 1 obtains a payoff close to c from playing τ^D. Since $c > 1$, it is optimal for player 1 to defect, hence we have $z(\bar{y}, k) = 0$.

2. $y_0 < \bar{y} = ak < y_1$. (Segment BC in figure 2) For $(\bar{y}, k)$ to be a fixed point, one should have $z(\bar{y}, k) = \bar{y}$ and $z(\bar{y}, k) \leq ah(\bar{y}, k)$. However, since $a < 1$, Lemma 1 (point **iii)**) implies $\mathbf{s}(\bar{y}, K) > 0$ hence $ah(\bar{y}, k) < \bar{y}$.

3. $\bar{y} = y_0$. (Segment CD in figure 2) For r_0 small enough, we have $ak \leq y_0 \leq Ak$ (by definition of y_0 and y_K), so we have $k_0 \leq k \leq \frac{A}{a} r^{-1/3}$. Then by Lemma 3, $\mathbf{w}(K, \underline{p})$ is close to 0 (because rk is very small), hence smaller than 1/3. For any $r \leq r_0$ with r_0 small enough, we have $K \leq r$ because $k \geq k_0 = r^{-1/2}$. So Lemma 3 implies $\underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p})) \leq d_0 r \mid \ln r \mid$, hence $\underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p})) \leq r^{1/2}$ for all $r \leq r_0$ if r_0 is small enough. Since $r \leq (\frac{A}{a})^3 k^{-3}$, we get $\underline{\mathbf{q}}(K, \mathbf{w}(K, \underline{p})) \leq (\frac{A}{a})^{3/2} k^{-3/2} << d_1 k^{-1}$ for r_0 small enough, which implies $z(\bar{y}, k) > y_K(\frac{d_1}{k}) \geq y_{K_0}(\frac{d_1}{k_0}) = y_0$.[47]

4. $y_0 < \bar{y} = y_0 + A(k - k_0) < y_1$. (Segment DA in figure 2). For $(\bar{y}, k)$ to be a fixed point, one should have $z(\bar{y}, k) = \bar{y}$ and $h(\bar{y}, k) \leq k$. However, since $A > \frac{\mu}{\mu - \lambda}$, we have $\bar{y} > y_K(\frac{d_1}{k})$,[48] and Proposition 5 implies $h(\bar{y}, k) > k$ (because otherwise there would exist a consistent and invariant belief such that $\underline{p} < \frac{d_1}{|\ln K|}$, contradicting Proposition 5) ■

Appendix E

We start with a Lemma used in both extensions.

Lemma 4 *There exists r_0 such that for any $r \leq r_0$, and for any belief $\bar{\beta}_1 = (p, \beta) \in [0,1] \times \Delta(\mathcal{T}_2)$, if $p \leq r^3$ then it is optimal for player 1 to switch to defect immediately.*

Proof. Assume that it is optimal for player 1 to cooperate and to wait for a lapse of time equal to t_0 without signal before defecting. Let $t_1 = \frac{2cr^2}{|b| + c}$, and assume $t_0 > t_1$. By waiting until t_0, player 1 obtains a payoff at most equal to $b(1 - e^{-rt_0})$ in the event player 2 has already switched to defect (this event has probability $1 - r^3$ at least), and at most equal to c in the event player 2 still cooperates. His expected payoff would thus be equal

[46] Recall that $\frac{\bar{y} - y}{\mu - K}$ is the time necessary to go from y to $\bar{y}$.

[47] This last inequality holds because $y_K(\theta)$ increases when K decreases and when θ decreases, and because, when k increases, both $K = e^{-k}$ and $\theta = \frac{d_1}{k}$ decrease.

[48] Indeed consider the function $f : k \to y_0 + A(k - k_0) - y_{e^{-k}}(\frac{d_1}{k})$. We have $f(k_0) = 0$ by definition of y_0, and it is easy to check that $f'(k) = A - \frac{\mu}{\mu - \lambda}(1 + \frac{1}{\ln k}) + O(e^{-k})$, hence $f'(k) > 0$ for all $k \geq k_0$, for k_0 large enough (that is, for r small enough).

to $b(1-e^{-rt_1})(1-r^3)+cr^3$ at most, hence negative by definition of t_1 for all $r \leq r_0$ if r_0 is small enough. Contradiction. So $t_0 \leq t_1$.

During t_0, the probability that either player receives a signal is at most equal to $2\mu t_0$. So with probability $1-2\mu t_0$, player 1 switches to defect at t_0. Besides, compared to the case where player 1 would have defected from the start, the behavior of player 2 after t_0 changes only if she was still cooperating *and* the realizations of the signal received by 2 is changed. So the behavior of player 2 is affected with probability $2\mu t_0 r^2$ at most. The expected gain from cooperating until t_0 compared to defecting from the start is therefore at most equal to

$$b(1-e^{-rt_0})(1-2\mu t_0)+2\mu t_0 r^2 c,$$

which is negative for all t_0 such that $0 < t_0 \leq t_1$. ■

Proof of Proposition 9:

Choose r_0, $r \leq r_0$ and $K, \underline{p}, B$ as in the proof of the main theorem. Define t^* as in the proof of Lemma 3 (part **ii)** a), bound on $w(K, \underline{p})$): $t^* = d_3 y_K(\underline{p})$ for some constant d_3 that only depends on λ, μ. For a consistent and invariant belief $\frac{3}{4} < \frac{y_K(p)}{|\ln K|} < 2\frac{\mu}{\mu-\lambda}$ (by Lemma 1 for the first inequality and Proposition 5 for the second one), and $|\ln K| \leq \frac{d_4}{cr}$ for some constant d_4 that only depends on λ, μ (otherwise we would have $v_1(\tau^D, \beta) > 1$, see Lemma 3). So $t^* = \frac{d_5}{cr}$ for some constant d_5 (that only depends on λ, μ.)

It was shown in the proof of Lemma 3 that after t^* periods of defection by player 1, player 2 still cooperates with probability at most equal to $e^{-2y_K(\underline{p})}$, hence with probability at most equal to r^3 (since $y_K(\underline{p}) > \frac{|\ln K|}{2}$ and since $K^{1/2} \leq r$ by the choice of r_0).

Consider the subgame where player 1 defected during t^*. Since player 1 received no (informative) signal during that time period, he believes that player 2 still cooperates with probability at most equal to r^3. By Lemma 4, it is optimal for him to keep on defecting. ■

Appendix F

Proof of Proposition 2:

The function ϕ_s just results from Bayesian updating (the fact that player 2 switches to defect at some rate K is not important because the jump occurs in an arbitrarily small amount of time). To get $\phi_{n,K}$, we use (2) with h very small and compute a first order approximation:

$$\phi_{n,K}(p,h)[1-(\mu p+\lambda(1-p))h]+\phi_s(p)[\mu p+\lambda(1-p)] = (1-Kh)p+o(h) \qquad (25)$$

Since $\phi_{n,K}(p,0) \equiv p$, dividing (25) by h and taking the limit $h \to 0$ shows that $\phi_{n,K}$ solves:

$$\frac{\partial \phi}{\partial h} = -\phi[K - (\lambda - \mu)(1 - \phi)]$$

Integrating that differential equation gives the result. ■

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., "On Failing to Cooperate when Monitoring is Private ", 1996, mimeo

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

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

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

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

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

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

[12] Mailath, G. and Morris, S., Repeated Games with Imperfect Private Monitoring: Notes on a Co-ordination Perspective, mimeo (1997)

[13] Mailath, G. and Samuelson, L., Your Reputation Is Who You're Not, Not Who You'd Like To Be, Working paper, University of Pennsylvania (1998)

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

[15] Sekiguchi, T. "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring", *Journal of Economic Theory*, **76**, 345-361 (1997)

[16] Schaefer, H., *Banach Lattices and Positive Operators*, Springer-Verlag, New-York 1974.

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

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