Moral Hazard and Private Monitoring

V. Bhaskar
Dept. of Economics
University of St. Andrews
St. Andrews
UK, KY16 9AL,
email: vb1@st-and.ac.uk

Eric van Damme*
CentER for Economic Research
P.O.Box 90153
5000 LE Tilburg
The Netherlands
email: Eric.vanDamme@kub.nl

This version: October 1997

Abstract

We analyze a model of repeated bilateral trade with moral hazard, where the quality of goods received can differ from the quality despatched due to deterioration during transportation. Since the sender does not observe the quality of good received and the receiver does not observe the quality despatched, we have a repeated game with with imperfect monitoring by private signals. The stage game has multiple Nash equilibria, which would allow cooperation in finitely repeated interaction. However, with private signals, the pure strategy equilibria of the twice-repeated game are degenerate, and cannot support any cooperation.We construct a mixed strategy equilibrium which supports partial cooperation. However this mixed strategy equilibrium cannot approximate the cooperative outcome even if the noise in the signals tends to zero. This failure of lower hemicontinuity in the sequential equilibrium correspondence is removed if we allow for extensive form correlation; i.e. we allow players to condition their second period actions upon a sunspot as well as the private signals. We use these ideas to show how efficient outcomes can be supported in infinitely repeated one-sided moral hazard.

Keywords: repeated games with imperfect monitoring, private signals, mixed strategies, sunspot equilibria.

JEL Categories: C73 (Stochastic and Dynamic Games), D82 (Asymmetric and Private Information), L15 (Information and Product Quality),

*This paper incorporates and supersedes an earlier paper by Bhaskar (1994), which contained most of the results of sections 2-4, and unpublished notes by van Damme on the role of mixed strategies. As compared to this earlier work , our new results, involving public randomization, are to be found in sections 5 and 6. We are grateful to Tilman Börgers, Dilip Mookherjee and Debraj Ray for useful discussions and also thank seminar audiences at Boston University, Essex, Econometric Society European Meeting (Maastricht), Institute of Economic Analysis (Barcelona), University College London, and the UK Game Theory Conference (Kenilworth). The first author thanks CentER for Economic Research (Tilburg) for its hospitality while some of this research was carried out.

1 Introduction

The theory of repeated games has been one of the most influential contributions of game theory to economics and other social sciences. In a single interaction, self-interested agents will generally indulge in opportunistic behavior. In repeated interaction, opportunism can be deterred, provided that agents are patient and are informed about the actions chosen by other agents in previous periods. This insight has diverse applications — for example, to the provision of product quality in a market economy, to collusion under oligopoly or as an explanation for "altruistic" behavior in social interaction.

The main results of this theory extend when individual agents' actions are not observed, provided that all agents observe a *public* signal which is informative of individual actions. Green and Porter (1984)[1] analyze the case of Cournot oligopoly with homogeneous products, where the output decisions of individual firms are unobservable, but the common market price is publicly observed. In equilibrium, all firms cooperate in the collusive agreement. Nevertheless, punishments are triggered after shocks which are sufficiently unfavorable, and hence agents incur payoff losses which may be attributable to the imperfectness of monitoring. However, these costs are small provided that the signals are sufficiently informative and players are patient, as is demonstrated by the "folk-theorem" for this class of games (see Fudenberg, Levine and Maskin (1994)). Finally, with public signals, as in the case of perfect monitoring, maximum cooperation is achieved by the use of severe punishments.

Rather less is known about repeated games where individual agents monitor other agents via *private* signals. An example of such a game was given by Stigler (1964) in his classic discussion of secret price-cutting in oligopoly. Each firm chooses price, and this price cannot be observed by other firms. The firm's sales are also privately observed, and are a noisy function of all prices. Apart from directly affecting its payoffs, each firm's sales is also a signal which is informative about the prices chosen by other firms. This signal may be very informative; however, the critical point is that both the firm's action

[1] Abreu, Pearce and Stachetti (1990) provide a general framework for the analysis for this class of games.

(its price) and its signal (its sales) are privately observed. With private signals, agents do not have common knowledge of whether cooperation is to continue or a punishment phase is to be started. This absence of common knowledge creates formidable problems. A general theory of repeated games with private signals has proved difficult to construct, since such games lack the recursive structure which was so fruitfully exploited in the case of public signals.

This paper analyzes repeated trade with moral hazard, which is an example of a simple repeated game with private signals. Our analysis highlights several new features which are specific to this class of games. First, we find that *cooperation requires defection* under private monitoring. This Orwellian feature is due to the fact that if cooperation is complete, each agent is unwilling to punish others in the event of observing a "bad" signal. Punishments can only be carried out if agents defect with some probability in equilibrium. Second, this need for defection imposes significant efficiency losses, which may be substantial, even if the monitoring technology is almost perfect. Finally, we show that irrelevant random public events (sunspots) may play an important role in sustaining efficiency, by allowing players to coordinate their behavior. Sunspots allow agents to forgive and forget, thereby reducing the severity of punishments. This plays a positive role in sustaining a high degree of cooperation.

We consider bilateral trade, where traders may supply each other a good of high quality or of low quality. Each trader's action (i.e. the quality supplied) is private information. Moreover, the quality of the good which is received by the recipient is also private. Traders also have a "no-trade" option, which implies that the one-shot trading game has multiple equilibria. This would normally allow the traders to support high quality trade in the initial periods in any finitely repeated interaction, if signals were public. We focus on the case where the trading game is repeated twice, and ask, can traders support the "efficient outcome" where they trade high quality in period one, and low quality in period two? Section 3 shows that such cooperation cannot be achieved by a pure strategy equilibrium, unless the signals received by the traders are highly correlated. Hence one must inevitably consider mixed strategy equilibria. In section 4 we construct an example of a mixed strategy equilibrium which supports the provision of high quality in period 1 with positive probability. Hence our first positive result is that we can achieve some cooperation. However, we find that we cannot approximate the efficient outcome even if the noise in the signals goes to zero — the sequential equilibrium outcome correspondence fails to be lower-hemicontinuous. In section 5 we allow players to observe the output of a public randomization device (a sunspot) at the end of period 1. This restores lower-hemicontinuity and allows players to approximate the best sequential equilibrium outcome of the noiseless game, as the noise in the signal goes to zero. Section 6 uses these ideas to show how cooperation can be sustained in the case of infinitely repeated one-sided moral hazard. The final section reviews the related literature and concludes.

2 The Basic Problem

Consider the following situation of bilateral trade with moral hazard. Two traders are exchanging goods of variable quality — to make things concrete, think of these as fruit, with one trader supplying lychees and the other mangoes. Each trader must independently make a preliminary investment, incurring a sunk cost F if they are to have the option to trade. If both traders pay this cost, they may proceed to trade. A trader can *cooperate* (i.e. choose action C) by sending fruit of high quality, or *defect* (i.e choose action D), by sending fruit of low quality. High quality fruit has value V_H to the recepient, which is greater than the value of low quality fruit, V_L. However, the cost of high quality to the supplier, C_H, also exceeds the cost of low quality, C_L. Payoffs as a function of the quality despatched by the trader (which we shall call the *action,* and is indicated by upper case letters) and the quality of fruit received (which we call the *signal,* and indicate using lower case letters) are shown in Fig.1. E is the no-trade option, and we have normalized payoffs by adding the sunk cost to each entry. (It is assumed that if only one partner decides not to trade (action E), the other trader loses the sunk cost but not any additional production cost.)

	c	d	e
C	$V_H - C_H$	$V_L - C_H$	0
D	$V_H - C_L$	$V_L - C_L$	0
E	F	F	F

Fig. 1

Assume that if one trader sends the other good fruit, there is a small probability, ϵ, that the fruit deteriorates en route, so that the latter receives low quality. Assume also that quality received by the two traders are independent events. The signalling technology is described by the stochastic matrix in Fig. 2.

	c	d	e
C	$1 - \epsilon$	ϵ	0
D	0	1	0
E	0	0	1

Fig.2

We may then write the payoffs to a trader as a function of his own action, and the action taken by the other trader, i.e the strategic form of this game, G, as follows

	C	D	E
C	$\tilde{V}_H - C_H$	$V_L - C_H$	0
D	$\tilde{V}_H - C_L$	$V_L - C_L$	0
E	F	F	F

The Game G

where $\tilde{V}_H = (1 - \epsilon)V_H + \epsilon V_L$ is the "expected quality" received when high quality is despatched. We assume that it is efficient to both traders to exchange high quality fruit, so that $\tilde{V}_H - C_H > V_L - C_L$.We shall assume that quality despatched and quality received are both unverifable, and hence high quality trade cannot be legally enforced. Clearly, the action C, of despatching high quality, is strictly dominated. However, both (D, D) and (E, E) are Nash equilibria of the game G, and there is also a mixed Nash equilibrium where each trader plays D with probability $\mu^* = \frac{F}{V_L - C_L}$ and E with probability $1 - \mu^*$. Since G has multiple Nash equilibria, this allows the traders to sustain cooperation if this trading game is repeated, even if only finitely many times. This is possible if low quality trade is sufficiently better than no-trade so that $V_L - C_L - F > \Delta C = C_H - C_L$. Assume this condition and focus attention on the case where G is played twice.[2]

Suppose that each player cannot observe the quality despatched by the other player, i.e. *actions* are unobserved. In line with the literature, but in contrast to the central focus of this paper, suppose that the quality *received* by any trader is commonly observed, i.e. the signals are *public.* This corresponds to a game with imperfect monitoring via public signals as in Abreu, Pearce and Stachetti (1990). Let each player adopt the following strategy: choose C in period one; in period two, play D if the signals are (c, c), and play E otherwise. To see that this strategy profile is an equilibrium, note that in period two, each player knows the action that his opponent will play for sure, and hence his own action is optimal at every information set. Given second period behavior, a deviation to D in period one is unprofitable. Equilibrium payoffs are given by

$$\tilde{V}_H - C_H + (1-\epsilon)^2(V_L - C_L) + (1 - (1 - \epsilon^2))F$$

This payoff is lower than the *efficient payoff* of $(\tilde{V}_H - C_H + V_L - C_L)$, which is an equilibrium payoff if the players actions were to be observed. (We call this the *efficient payoff* since this is the maximum payoff that each player can achieve in any equilibrium). Imperfect monitoring via public signals creates an inefficiency relative to the efficient payoff, but this inefficiency is of order ϵ, and vanishes as ϵ tends to zero.

Consider now an alternative information structure which is the focus of this paper, where each trader observes the quality of fruit he receives but does not observe the quality received by the other trader — signals are *private.* Hence neither the quality sent nor the quality received by trader i are mutual knowledge between the traders, although they

[2] If trade is seasonal, as is likely in the fruit example, the finitely repeated game may be a better representation of interaction than an infinitely repeated game. In addition, the two-period example allows us to characterize the efficiency properties of *all* equilibria.

could be arbitrarily close to being so if ϵ is small. This lack of mutual knowledge creates a dramatic discontinuity — we cannot support the playing of C in period one in any pure strategy equilibrium. Suppose that C is chosen by both traders in period one. This can only be optimal for each trader if he believes that the other trader will reward signal c and punish signal d. Hence each player's strategy must be of the type: play C in period 1; in period two, play D on receiving signal c, and play E on receiving signal d. [3]However, such a strategy is not a best response to itself; it is not optimal for a trader who receives signal d to carry out this punishment. Suppose that I am a player who believes that my opponent is playing such a strategy. If I observe the signal d, I should attribute this to the error in the signalling technology — the application of Bayes' rule to my opponent's strategy implies that this is the only event which has positive probability. Since I have chosen C in period one, I know that my opponent will receive signal c with very high probability, $1-\epsilon$. Hence it is optimal for me to continue with D, and ignore the signal I have received. Since varying second period behavior with the first period signal is not optimal, this makes it impossible to support the playing of C with probability one in the first period.[4]

Private signals create a *coordination problem*. Cooperative equilibrium requires that the players vary second period behavior, by playing the good equilibrium when they both receive a good signal, and the bad equilibrium when either signal is bad. With independent signals, these switches across stage game equilibria are uncoordinated, and hence inconsistent with optimality in the second period. This suggests a natural resolution to the problem – if a trader who receives a bad fruit can get on to the telephone and complain to the other, this public communication can generate a public signal which may be able to resolve the coordination problem. The role of communication in generating the requisite public signals and thereby ensuring coordination has been explored by Compte

[3]A player could also punish by playing the mixed equilibrium, but the argument which follows also applies in this case.

[4]This argument appears to be quite general — given independent signals where every signal has positive probability under any action profile, and generic payoffs in the stage game, the pure strategy equilibria of the twice repeated game must be degenerate, i.e. repetitions of stage-game Nash equilibria. By using induction, this result may also be extended to any number of finite repetitions.

(1994) and Kandori & Matsushima (1994). These authors show that with communication, one has a folk theorem with pure strategies for infinitely repeated games with private signals. The question therefore arises, can communication help support cooperation in our context of twice-repeated bilateral trade? The answer to this question is unfortunately in the negative.Consider a tentative equilibrium in which the players play C in period one, and condition their second period behavior upon whether each of them "complains" at the end of this period. Such communication can clearly solve the coordination problem, since it is common knowledge between the players. However, neither player will have an incentive to complain, even if he observes signal d. To ensure that playing C is optimal, the traders must both play D in period two if neither complains, and further, if trader 1 complains and trader 2 does not complain, trader 2 must be punished. To punish trader 2, the players must play either the (E, E) equilibrium or the mixed strategy Nash equilibrium of the stage game, so that trader 1 punishes himself if he complains. By using Bayes rule, trader 1 deduces that trader 1 has indeed played C, and the signal d is due to the noise. Since he himself has played C in period one, he believes that trader 2 has received the signal c with high probability and is very likely not to complain, and to continue to play D if he does not receive a complaint. Consequently, it is optimal for trader 1 not to complain, irrespective of the signal he receives, which makes the playing of C in period one impossible to support.[5]

This argument extends if the trading game is finitely repeated for any number of periods. The above argument implies that at the end of the penultimate period, traders have no incentive to complain. This implies that traders will not play C in the penultimate period, and will also not have incentive to complain in at the end of the period prior to the penultimate period. By backwards induction, one can show that no trader will ever complain, and hence communication cannot support cooperation. We turn therefore to a

[5]Compte and Kandori & Matsushima allow the players to condition their actions upon the communications as well as a public randomization device. It is easy to check that allowing a public randomization device does not help in our example. With such a device, players can play a convex combination of Nash equilibria of G in period two after any pair of exchanged messages. This allows players to achieve any stage game payoff pair in the line segment joining $(V_L - C_L, V_L - C_L)$ and (F, F). Once again, players have common interests – it is impossible to punish trader 2 without simultaneously punishing trader 1.

more complete non-cooperative analysis without communication.

3 Analysis

We analyze the following stage game. Each player $i \in \{1,2\}$ simultaneously chooses an action $a_i \in A_i$, where $A_1 = A_2 = \{C, D, E\}$, and $A = A_1 \times A_2$. Let $\Omega_1 = \Omega_2 = \{c, d, e\}$ be the set of possible signals which may received be received by players 1 and 2. Given $a \in A$, nature selects $\omega = (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$, with probability $q(\omega, a)$. Player i is informed of the realization of ω_i and receives a payoff $u_i(a_i, \omega_i)$. Let $p(\omega_1, a) = \sum_{\omega_2 \in \Omega_2} q((\omega_1, \omega_2), a)$ denote the marginal distribution of ω_1 given a, and let $p(\omega_2, a)$ denote the marginal distribution of ω_2, defined similarly. Hence player $i's$ payoff from an action profile $a = (a_1, a_2)$ is

$$v_i(a_1, a_2) = \sum_{\omega_i \in \Omega_i} u_i(a_i, \omega_i) p(\omega_i, a) \tag{1}$$

Hence the strategic form of the stage game is defined by the strategy sets A_1 and A_2, and the payoff functions v_1 and v_2.

We focus on the following signalling technology, where signals are private, but could possibly be correlated. Such correlation could arise due to, for example, correlated weather shocks which affect the fruit that both players receive. The distribution of signals conditional on $a = (C, C)$ is given by

	Trader 2's signal		
		c	d
Trader 1's signal	c	$(1-\epsilon)^2 + \rho\epsilon(1-\epsilon)$	$(1-\rho)\epsilon(1-\epsilon)$
	d	$(1-\rho)\epsilon(1-\epsilon)$	$\epsilon^2 + \rho\epsilon(1-\epsilon)$

Fig. 3 Distribution of signals conditional on (C, C)

If $a = (C, D)$, $\omega = (d, c)$ with probability $1 - \epsilon$, and $\omega = (d, d)$ with probability ϵ. If $a = (D, C)$, $\omega = (c, d)$ with probability $1 - \epsilon$, and $\omega = (d, d)$ with probability ϵ. If $a = (D, D), \omega = (d, d)$ with probability one. This signalling structure is parametrized by ϵ and ρ, where ϵ is the level of "noise", and ρ is the degree of correlation between signals.

$\rho = 0$ corresponds to case where the signals are independent, while $\rho > 0$ correponds to positive correlation between signaling errors. If $\rho = 1$, the signals are perfectly correlated, and this is equivalent to the situation where signals are public rather than private. We shall assume henceforth that $\rho < 1$.

Our focus is on the twice repeated game, which we denote $G^2(\epsilon, \rho)$.Players maximize the sum of expected payoffs in the two stages. A pure strategy for a player i in $G^2(\epsilon, \rho)$ is a pair $s_i = (f_i, g_i)$ where $f_i \in A_i$ is the action taken in the first period, and $g_i : \Omega_i \times A_i \to A_i$ specifies the action taken in the second period as a function of the player's first period action and the signal he receives. Let S_i be the set of pure strategies for player i. A pair of pure strategies $s = (s_1, s_2)$ generates a repeated game payoff $V_i(s)$ as follows

$$V_i(s) := v_i(f_1, f_2) + \sum_{\omega_2 \in \Omega_2} \sum_{\omega_1 \in \Omega_1} v_i[g_1(f_1, \omega_1), g_2(f_2, \omega_2)] p(\omega_1, (f_1, f_2)) p(\omega_2, (f_1, f_2)) \quad (2)$$

A mixed strategy for a player i is a probability vector σ_i, where $\sigma_i(s_i)$ denotes the probability assigned to the pure strategy s_i. Let Σ_i be the set of mixed strategies for player i, and let $\Sigma = \Sigma_1 \times \Sigma_2$ be the set of mixed strategy profiles. We extend the payoff function V_i to Σ in the usual way: given a mixed strategy profile $\sigma = (\sigma_1, \sigma_2)$, we have

$$V_i(\sigma) = \sum_{(s_1, s_2) \in S} V_i(s_1, s_2) \sigma_1(s_1) \sigma_2(s_2) \quad (3)$$

Given $\sigma = (\sigma_1, \sigma_2)$, let $\sigma|\sigma_i'$ denote the profile derived by replacing σ_i with σ_i'. The profile $\sigma = (\sigma_1, \sigma_2)$ is a Nash equilibrium if for $i = 1, 2$, $V_i(\sigma) \geq V_i(\sigma|\sigma_i') \forall \sigma_i' \in \Sigma_i$.

The main arguments of this paper pertain to Nash equilibria.[6] However, it is usual in dynamic games to require that equilibrium be sequential, so that players are behaving optimally at each information set, including those that are not reached under the strategy profile. To define this, consider player $i's$ beliefs about player $j's$ second period action at the information set (ω_i, a_i) (i.e. when he has played a_i and received signal ω_i). Player $i's$

[6] Our focus is on strategy profiles where C is played in period one, and in this case signals c and d will both be observed with positive probability. The Nash equilibrium criterion, which requires optimal behavior at all information sets which are reached, is hence sufficient.

beliefs about $j's$ action are derived from this information and his prior knowledge of $j's$ strategy, σ_j. If the signal ω_i has positive probability under σ_j, $i's$ beliefs are given by an application of Bayes rule. If ω_i does not have positive probability under σ_j, then i infers that j has deviated from σ_j, and furthermore, given our signalling technology, i can also infer $j's$ action. Since σ_j is defined for all first period actions of player j, this suffices to ensure that $i's$ beliefs are uniquely defined in this case as well.

It will be useful to set out more explicitly the conditions that any sequential equilibium must satisfy. Let $\sigma = (\sigma_1, \sigma_2)$ be an equilibrium, and let $s_i = (f_i, g_i)$ be a pure strategy which is the support of σ_i, where $i \in \{1, 2\}$. Consider second period behavior, at an information set (a_i, ω_i). Since C is a strictly dominated action, it cannot be played by any player at any information set. Let $\mu_i(\omega_i, a_i; \sigma_j)$ be player $i's$ *belief* – the probability assigned by i to event that j will play D in period two. Hence i believes that j will play E with complementary probability, $1 - \mu_i(\omega_i, a_i; \sigma_j)$.This must satisfy

$$g_i(a_i, \omega_i) = D \Rightarrow \mu_i(\omega_i, a_i; \sigma_j) \geq \mu^* \tag{4}$$

$$g_i(a_i, \omega_i) = E \Rightarrow \mu_i(\omega_i, a_i; \sigma_j) \leq \mu^* \tag{5}$$

Finally, first period behavior must also be optimal, i.e. it must be optimal to play the prescribed action f_i in period one, rather than deviate to any alternative action:

$$V_i((f_i, g_i), \sigma_j) \geq V_i((a_i, g_i), \sigma_j) \forall a_i \in A_i \tag{6}$$

We are now in a position to discuss the sustainability of cooperation via a pure strategy equilibrium, with correlated signals. With correlated signaling errors, if a player receives a bad signal, this makes it more likely that his opponent has also received a bad signal. Consequently an agreement to punish on receiving a bad signal could be made self enforcing. However, the degree of correlation must be large enough. Define the strategy α as follows

Strategy α: 1st period: C. 2nd period:D if (C, c), E otherwise.

Consider the sustainability of the strategy profile (α, α).[7] To check that this is a Nash equilibrium we need to see that second period behavior is optimal. If my opponent is playing the strategy α, then he will play D in period 2 if he has observed the signal c, and will play E if he has observed d. For the second period behavior dictated by α to be optimal, we must have

$$\mu_i(c, C; \alpha) = (1 - \epsilon) + \rho\epsilon \geq \mu^*. \tag{7}$$

$$\mu_i(d, C; \alpha) = (1 - \rho)(1 - \epsilon) \leq \mu^* \tag{8}$$

In addition, it must be optimal to play C in period one, rather than deviating by playing D in period one and E in period two. Let $\Delta C = C_H - C_L$— this is first period gain to deviating by producing low quality. This must be less than the second period loss from deviation, i.e.

$$\Delta C \leq [((1 - \epsilon)^2 + \rho\epsilon(1 - \epsilon))(V_L - C_L) + \epsilon F] - F \tag{9}$$

Let $\rho \geq 0$, so that ϵ and ρ both lie in the unit interval. Inequalities (7-9) are graphed in Fig. 4. The shaded area in this figure shows values of ϵ and ρ such that these inequalities are satisfied, and (α, α) is an equilibrium. The key features of this figure are summarized in the following proposition.

Proposition 1 *i)If $\rho \geq 1-\mu^*$, cooperation can be supported by a pure strategy equilibrium if ϵ is sufficiently small.*

ii)If $\rho < 1 - \mu^*$, cooperation cannot be supported by a pure strategy equilibrium if ϵ is sufficiently small.

iii)If ρ is close to but less than $1 - \mu^*$, cooperation can be supported if ϵ is neither too large nor too small.

[7] Any pure strategy equilibrium where C is played in period one must be similar to (α, α), since signal c must be rewarded and d must be punished. What happens after signal e is irrelevant.

Observe that the negative pure strategy result is robust, correlation must be sufficiently high for cooperation to be supported. Most intriguing is part (iii) of the proposition, on the relation between the level of noise, ϵ, and cooperation at intermediate levels of correlation. (8) will not hold if ϵ is small and close to zero, but Fig. 4 shows that this inequality can be satisfied for larger values of ϵ. However, ϵ must not be too large since otherwise (9) will not be satisfied. Hence the set of pure strategy equilibrium outcomes is not monotone in ϵ.This contrasts sharply with Kandori's (1992) result, that with imperfect monitoring by public signals, the pure strategy equilibrium set is a monotonically decreasing function of noise.[8]

Finally, we note that conflict between Pareto-dominance and risk dominance in the stage game, may facilitate cooperation in the repeated game. Consider the effect of changes in μ^* : for small values of ϵ, the critical inequality is clearly (8), which is easier to satisfy if μ^* is larger. In particular, if μ^* is larger, less correlation is required in to satisfy (8), and to support cooperation in the repeated game. However, μ^* is simply the "basin of attraction" of the inefficient equilibrium. As μ^* increases, the relative riskiness of the two stage game equilibria changes — (E, E) becomes less risky as compared to (D, D).

We shall henceforth focus attention upon the case where $\rho < 1 - \mu^*$, when cooperation cannot be sustained via pure strategies. We refer the interested reader to Mailath and Morris (1997), who discuss correlated signals in greater detail. They relate correlated signals to the literature on approximate common knowledge, and also prove a folk theorem for infinitely repeated games with private signals if these signals are sufficiently highly correlated.

4 Mixed Strategies: Their Role & Limitations

We now construct a mixed strategy equilibrium which allows us to support partial cooperation in the twice repeated game for any level of correlation between signals. In order to understand the role of mixed strategies, it is useful to interpret the reason why pure

[8] This paradoxical finding also applies with mixed strategies, as proposition 3 below shows.

strategies are unable to support any cooperation. Focus on the case where signals are independent so that $\rho = 0$. Observe that in this case, from (7) and (8) that $\mu_i(c, C, \alpha) = \mu_i(d, C, \alpha) = 1 - \epsilon$. In other words, if a player "knows" his opponent's strategy (as is implicit in a pure strategy equilibrium), his beliefs regarding his opponent's action in period two depend only upon his prior knowledge, and are insensitive to the signal he receives. To make a player willing to respond to the signal, we must ensure that it conveys some information about his opponent's second period actions in equilibrium. More specifically, a player will be willing to respond differently to different signals only if these signals indicate that his opponent is likely to play differently.[9]This is possible if we allow for mixed strategies, since the player's prior beliefs will not be degenerate, and the signal allows him to learn which pure strategy his opponent is playing.

Consider the following pure strategies for the repeated game.

Strategy α: 1st period: C. 2nd period:D if $(\omega_i, a_i) = (c, C)$, E otherwise.

Strategy β: 1st period: D. 2nd period:E.

The payoff matrix for these two supergame strategies is:

	α	β
α	$\tilde{V}_H - C_H + V_L - C_L - O(\epsilon)$	$V_L - C_H + F$
β	$\tilde{V}_H - C_L + F$	$V_L - C_L + F$

where $O(\epsilon)$ is a term of order ϵ, $O(\epsilon) = [(1-\epsilon)^2 + \rho\epsilon(1-\epsilon)](V_L - C_L) + \epsilon F$.

Confining attention to the pure strategy set $\{\alpha, \beta\}$ for each player, we see that α is a strict best response to α if ϵ is sufficiently small and β is a strict best response to β. Hence the above payoff matrix also has a symmetric mixed strategy equilibrium where each player plays α with probability π and β with probability $1 - \pi$, where $\pi = \frac{\Delta C}{V_L - C_L - F - O(\epsilon)}$. Call this mixed strategy σ.We now show that the symmetric strategy profile (σ, σ) is a an equilibrium of the repeated game.

[9]Alternatively, a player can be made willing respond to the signal even with constant beliefs if $\mu^* = 1 - \epsilon$, so that he is indifferent between his two actions and takes different actions at different information sets. We discuss this possibility, due to Kandori (1991) after proposition 2.

Proposition 2 *The symmetric strategy profile where each player plays* σ *is an equilibrium of* $G^2(\epsilon, \rho)$ *for any* $\rho < 1$ *if* ϵ *is sufficiently small.*

Proof: Assume that the opponent plays σ. It is easily seen that any strategy that starts by playing E is strictly inferior. Write $\mu_i(.;\sigma)$ for the beliefs induced by σ, i.e. the probability that the opponent will play D at $t = 2$. Then

$$\mu_i(c, C; \sigma) \to 1 \text{ as } \epsilon \to 0 \tag{10}$$

$$\mu_i(d, C; \sigma) \to 0 \text{ as } \epsilon \to 0 \tag{11}$$

$$\mu_i(., D; \sigma) = 0 \tag{12}$$

At information set (c, C), I know that my opponent has played α and that he received signal c with probability almost 1, and hence (10) follows. At information set (d, C), Bayes rule implies

$$\mu_i(d, C; \sigma) = \frac{\pi\epsilon}{(1-\pi) + \pi\epsilon} \times \frac{(1-\rho)\epsilon(1-\epsilon)}{\epsilon} \tag{13}$$

so (11) follows from the fact that $0 < \rho < 1$ and $0 < \pi < 1$. Finally, (12) follows since the opponent is sure to receive signal d after D and since both α and β play E after d. (10-12) together with the fact that both (D, D) and (E, E) are strict equilibria of G imply that for ϵ small enough, D is the unique best response at (c, C), and E is the unique best response at other information sets at $t = 2$. It follows that both α and β prescribe best responses to σ at $t = 2$. Since, by construction, α and β are also best responses at $t = 1$, (σ, σ) is an equilibrium of the game. ■

We are not the first to show that partial cooperation can be sustained in a repeated game with private signals. Kandori (1991) shows this for the twice repetition of a stage game that has a unique mixed strategy equilibrium, where the private signals are independent. Kandori's example does not work if the signals are correlated to any extent (correlation is not a problem in our case). Kandori's equilibrium requires history-dependent randomization, i.e. a player is indifferent between two actions in period two and he chooses one or the other depending upon which history materializes. As Bhaskar (1996, 1998) shows more generally, such mixed equilibria cannot be purified in the sense of Harsanyi (1973), i.e. they are not robust and will vanish once slight incomplete information about payoffs is introduced (In that case, a player will almost always behave in the same way after different histories, since he strictly prefers one action above the other). Note that in the equilibrium constructed above, a player is required to randomize only at stage 1 and has strict incentives to follow the recommendations of his strategy at stage 2.[10] Hence the equilibrium does not require history-dependent randomization, and it is not difficult to construct equilibria of incomplete information games that approximate it. In other words, ours appears to be the first *robust* equilibrium that sustains partial cooperation in the case of imperfect private monitoring. Furthermore, the idea of the construction presented here appears to be generalizable to other contexts. In subsequent work, Sekiguchi (1996) has constructed a cooperative mixed strategy equilibrium of the infinitely repeated prisoners' dilemma which uses a similar idea. We discuss Sekiguchi's work in greater detail in the concluding section.

Although the mixed strategy equilibrium supports partial cooperation, the probability with which the players play C in period one is bounded away from one even if ϵ is arbitrarily small. To see this, observe that π, the probability with which the strategy α is played, tends to $\frac{\Delta C}{V_L - C_L - F} < 1$ as $\epsilon \to 0$. Hence the equilibrium payoff in the game without any noise cannot be approximated by this mixed equilibrium, in contrast with the situation where signals are publicly observed. We now show that this result holds more generally — the cooperative equilibrium under perfect monitoring cannot be approximated under imperfect monitoring even if the noise in the signals goes to zero. In other words, the

[10] It may also be useful to note that the mixed strategies we consider generate correlated signals endogenously, and indeed approximate common knowledge of these signals. See Mailath and Morris (1997) for a discussion.

sequential equilibrium outcome correspondence is not lower-hemicontinuous.[11]

Proposition 3 *If $\rho < 1-\mu^*$, the efficient outcome where both traders produce high quality in period one, and low quality in period two cannot be approximated by any equilibrium of $G^2(\epsilon,\rho)$, as $\epsilon \to 0$.*

Proof: See Appendix. ■

The basic idea of the proof this proposition is as follows. In any approximately efficient mixed strategy equilibrium, both players must play *good* pure strategies with high probability, where a good strategy is defined as one which plays C in the first period, and responds to the signal c with D. If both players are playing good strategies with high probability, then any strategy which plays D in the first period will have a higher first period payoff. Hence equilibrium requires that the signal d be punished, i.e. both players must attach positive probability to a good strategy which punishes signal d by playing E. However, such a punishing good strategy plays differently after signals c and d, and for this to be optimal for player i, the signal d must indicate that player j is likely to play E. If $\rho < 1-\mu^*$ and ϵ is small, this implies that j must attach positive probability to a *bad* strategy, which is defined as one which plays D in period one, and responds to signal c with E. However, a bad strategy can earn at most a payoff of $\tilde{V}_H - C_L + F$, which is strictly less than the efficient payoff. Consequently, equilibrium payoffs are bounded away from efficiency.

5 Sunspots & Efficiency

The inefficiency result of the previous section arises for the following reason: players must assign positive probability to bad strategies, which defect in equilibrium. However, in order to induce players to cooperate, defectors must also be punished. The problem is

[11] This failure of lower-hemicontinuity is with respect to the information structure, and hence quite different from the example of Radner, Maskin and Myerson (1986). This example considers the behavior of equilibrium payoffs in a repeated game with public signals as the discount rate varies, with a fixed information structure.

that the punishment for defection is so severe that this reduces the payoff to defection, and hence equilibrium payoffs. This intuition suggests that if one can soften the punishment of defector, this may help reduce the inefficiency, by raising the payoff to a defector, thereby increasing equilibrium payoffs.

How do we soften the punishment to defection? One possibility is that in period two, each player does not always punish the signal d, but merely punishes with some probability, by randomizing between E and D in the event of receiving signal d. However, such randomization at the individual level is infeasible, since each player has strict incentives to play D at this information set. What is required is that player can agree to forget past transgressions in a *coordinated* way. A sunspot, i.e. the realization of a commonly observed random variable, can play this role, even though this sunspot is uninformative about players' past actions. Intuitively, players can agree to forget about past transgressions on sunny days, reserving punishments for defections only for the days where the weather frowns. By doing so they can ensure that defectors are deterred, but not too harshly. Formally, the sunspot allows for *extensive-form correlation,* which transforms the base game by convexifying the set of equilibrium payoffs, allowing the two players to achieve any payoff in the interval $[F, V_L - C_L]$. Consequently, a player who chooses D in period one can be punished so that her payoff loss in period 2 is arbitrarily close to her payoff gain in period 1. Since there is no overall payoff loss from playing a bad strategy, this enables both players to play a bad strategy with small probability.

Assume that at the beginning of period two, before the players choose their actions, players can publicly observe the outcome ϕ of a random variable Φ, which is uniformly distributed on $[0, 1]$. The sunspot convexifies the set of equilibrium payoffs of G. Specifically, for any $m \in [0, 1]$, the correlated strategy $z = (z_1, z_2)$ with

$$z_1(\phi) = z_2(\phi) = \begin{cases} D \text{ if } \phi \leq m \\ E \text{ if } \phi > m \end{cases} \tag{14}$$

is a correlated equilibrium of G. By varying m, any payoff Z in $[F, V_L - C_L]$ can be obtained in this way. Note that such a correlated equilibrium z is strict: if a player believes that his opponent plays z_j with probability greater than $\min\{\mu^*, 1-\mu^*\}$, then it is optimal to play z_i himself. Let z be such a correlated equilibrium of G with payoff Z, and modify the strategies α and β from the previous section such that E is replaced by z. The only thing that changes in the payoff matrix is that F has to be replaced by Z. Provided that

$$Z + \Delta C < V_L - C_L - O(\epsilon) \tag{15}$$

(an inequality which is satisfied for Z sufficiently close to F), (α, α) and (β, β) are still strict equilibria of this 2×2 game, and as in the previous section, there exists a mixed strategy equilibrium of this payoff matrix, where α is played with probability π and β with probability $1 - \pi$. The claim of the previous section, that this is an equilibrium of the repeated game, continues to apply. Observe from the proof that the only essential change occurs when considering the information set (C, d). For any given π, I attach a probability greater than $\min\{\mu^*, 1-\mu^*\}$ to my opponent continuing with z_j provided that ϵ is sufficiently small.[12] Since z is strict, it is optimal for me to continue with z_i as well. Now, investigate the consequences of varying Z. By increasing Z to the upper bound from (15), (while at the same time reducing the noise level accordingly), the probability π from can be increased to 1. In other words, players will play (α, α) with probability close to one, and will obtain a payoff close to the efficient one.

We have therefore proved.

Proposition 4 *If players can observe a common sunspot in addition to their private signal, then they will be able to approximate the efficient payoff when the noise, ϵ, is small.*

Observe that time at which the output of the public randomization device is observed by both players is crucial. This must be *after* players have chosen their actions in period 1, but before they choose actions in period 2. In other words, *extensive form correlation* is essential. Extensive-form correlation was introduced by Myerson (1986), who also pointed

[12]The relevant condition is inequality (13), which specifies how small ϵ must be given π.

out this allows greater strategic possibilities than normal-form correlation.[13]

6 Infinitely Repeated Interaction

We now the two main ideas developed in the earlier sections, of using mixed strategies, and conditioning behavior on sunspots, may fruitfully be employed in the instance of infinitely repeated moral hazard. We focus on the interaction between a customer and a seller, and focus on one-sided moral hazard, on the part of the seller.

In each period the two parties choose their actions simultaneously. The customer must decide whether to buy (action B) or not buy (action N). If the customer chooses N, both parties get a payoff of zero. The seller must choose between supplying high quality (action C) and low quality (action D). If the seller chooses C, the customer receives high quality (signal c) in the event that he buys with probability $(1-\epsilon)$ and low quality with probability ϵ. If the customer chooses D, the quality received (signal d) is always low. The customer's valuations for the two types of products are V_H and V_L, and let $\tilde{V}_H = (1-\epsilon)V_H + \epsilon V_L$ be the expected value to the customer when the seller chooses C. We assume that the price the customer pays the seller, P, is fixed. By proving a positive result with a fixed price, without P being a strategic variable, we ensure that any division of the surplus between customer and seller can be supported as an equilibrium. The strategic form of the game Γ is given by

		SELLER	
		C	D
CUSTOMER	B	$\tilde{V}_H - P, P - C_H$	$V_L - P, P - C_L$
	N	$0, 0$	$0, 0$

Γ

where $\tilde{V}_H > P > V_L, P > C_H > C_L$.

[13]Extensive form correlation via a public randomization device has also been employed in a recent paper by Harris, Reny and Robson (1995) in order to restore existence of subgame perfect equilibrium infinite games of almost perfect information.

At the end of the period, the customer's action is observed by both players, so that the seller observes whether B or N has been chosen. The customer observes the signal regarding the seller's choice if and only if he buys, and this signal is private. An alternative interpretation of our model is of worker-employer interaction in an efficiency wage model as in Shapiro and Stiglitz (1984) , where the employer (=customer) observes a private signal of the effort made by the worker (=seller), and provides effort incentives by the threat to fire. In essence our focus is on the credibility of this threat when the worker cannot observe the employer's observation of his effort.

This game is repeated infinitely often, and players discount payoffs at a common rate δ. Normalize the discounted stream of payoffs by multiplying these by $(1-\delta)$.We look for a "cooperative" equilibrium of this infinitely repeated game where the customer buys and the seller produces high quality. Clearly, the equilibrium cannot be of the usual trigger type where the seller always provides good quality and the customer always buys as long as he receives good quality. For, if the seller uses such a pure strategy, then bad quality is exclusively due to bad luck, and it would be optimal for the customer to continue to buy also after receiving bad quality. In other words, to give the buyer an incentive to punish (to stop buying after receiving bad quality), the seller must randomize. For the seller to randomize, she must be indifferent between supplying low and high quality. Now, supplying low quality yields only a modest one-period gain, while supplying high quality yields a long stream of positive payoffs if ϵ is small. It follows that the seller will be indifferent only if she is sufficiently impatient. In other words, if δ is sufficiently large, also a mixed trigger strategy equilibrium will not exist, as the seller will prefer to supply high quality for sure if the buyer punishes low quality.

At this stage, sunspots can be used to provide the proper incentives. Players can shorten the length of the game by terminating it when a certain sunspot occurs, and re-starting a new game. Formally, let players observe a random variable that is uniformly distributed on $[0,1]$, and let them terminate the game the first time the random variable takes a value less than m. Effectively, players then have a discount factor of δm, and by choosing m appropriately we can make it in the seller's interest to randomize between high and low quality. The occurrence of the sunspot (a realization less than m) also allows the players to coordinate, to forget past events and to start the game anew after it occurred. In this way it is guaranteed that punishments are not too severe.

We now formalize the above arguments. The strategies for the two players in a single game of stochastic length are represented most simply by the two automata in Fig. 5. The customer plays a pure strategy; she begins by buying, and continues to do so as long as she receives signal c. She switches to N if she ever receives signal d, and never buys again. The seller begins in the cooperative phase, where she randomizes between actions C and D. If her realized action is C and the customer has bought, she continues in the cooperative phase. If her realized action is D or the customer fails to buy, she switches to the punishment phase where she plays D thereafter. [14]

We show that this strategy profile is an equilibrium of the repeated game. We shall evaluate payoff to any strategy in a single game of stochastic length — i.e. we disregard the payoffs which accrue when a new game begins. This is permissible since the payoffs in any new game do not depend upon current actions.

Consider the seller's payoffs in the cooperative phase. By playing D, the seller earns $(P - C_L)$ in the current period, and zero in the remainder of the game of stochastic length. Hence her expected payoff from D is $(1-\delta m)(P-C_L)$. The payoff from playing C can be written as the expected payoff from playing C in the current period, and playing D in the next period. Since the customer will buy in the next period with probability $1-\epsilon$ (given that the seller is playing C currently), this equals

$$(1-\delta m)(P-C_H) + \delta m[(1-\epsilon)(1-\delta m)(P-C_L) + \delta m \epsilon 0] \tag{16}$$

Hence the payoffs to actions C and D will be equal if

$$m = \frac{\Delta C}{\delta (P - C_L)(1-\epsilon)} \tag{17}$$

[14] This specifies the two players' actions at all information sets which are reached. In the appendix, we provide a complete description of players' strategies, by specifying their actions at unreached information sets.

which is easy to ensure by an appropriate choice of $m \in (0,1)$, provided that $\delta > \frac{\Delta C}{(P-C_L)(1-\epsilon)}$, which is less than 1. Hence by public randomization, one can ensure that the seller is willing to randomize in the cooperative phase.

Consider now the conditions that must be satisfied by π, the seller's randomization probability in the cooperative phase, so that the customer's trigger strategy is optimal. Let $\tilde{\pi} = \pi(1-\epsilon)$. Suppose that we are either in period one, or in any subsequent period such that the customer has always bought in the past, and has always observed high quality c. If the customer fails to buy at this point, his continuation payoff is 0. Hence his continuation payoff from the recommended action (B), v_c, must satisfy

$$v_c = \tilde{\pi}[(V_H - P)(1-\delta m) + \delta m v_c] + (1-\tilde{\pi})(1-\delta m)(V_L - P) \geq 0 \tag{18}$$

$$v_c = \frac{(1-\delta m)[\tilde{\pi}(V_H - P) + (1-\tilde{\pi})(V_L - P)]}{1 - \delta m \tilde{\pi}} \geq 0 \tag{19}$$

This inequality is equivalent to the condition that $\tilde{\pi} > \frac{P-V_L}{V_H-V_L}$ and can hence satisfied for some $\pi \in (0,1)$ if ϵ is small enough.

It remains now to show that the customer is willing to switch from the cooperative phase to the punishment phase, on observing a bad signal. Assume therefore that the consumer's information about the past is of the form $(B,c), (B,c), ..., (B,c), (B,d)$, i.e. he has always bought in every previous period, and where quality has been high in every period except the last. At this point the customer believes that the seller will choose high quality in the next period with probability

$$\theta(\pi) = \frac{\pi\epsilon}{\pi\epsilon + (1-\pi)} \times \pi \tag{20}$$

Note that $\theta(\pi)$ is strictly less than π. Let $\tilde{\theta}(\pi) = \theta(\pi)(1-\epsilon)$. The customer's continuation payoff from complying with the trigger strategy is clearly zero. His one period deviation from this strategy is to play B in the current period, and to condition his state on the signal he receives — he continues in the cooperative phase if this signal is c, and in the punishment phase if the signal is d. (This follows from the description of his continuation strategy set out in the appendix). His payoff from this one-period deviation is a similar expression to (18), with $\tilde{\theta}(\pi)$ replacing $\tilde{\pi}$, and must satisfy

$$\tilde{\theta}(\pi)[(V_H - P)(1-\delta m) + \delta m v_c] + (1-\tilde{\theta}(\pi))(1-\delta m)(V_L - P) \leq 0 \qquad (21)$$

Since $\theta(\pi) < \pi$, the left hand-side of (21) is strictly less than the left-hand side of (18), one can make the former negative and the latter positive by an appropriate choice of π. (Note that both inequalities are increasing in π, and strictly positive for π sufficiently close to one, and strictly negative for π sufficiently close to zero). Hence we have constructed an equilibrium of the repeated game with partial cooperation.

We now show that can approximate the cooperative outcome arbitrarily closely provided that the noise tends to zero. Let $\pi^*(\epsilon)$ be the maximum value of π such that (21) is satisfied. Observe that $\theta(\pi) \to 0$ as $\epsilon \to 0$. Consequently, $\pi^*(\epsilon) \to 1$ as $\epsilon \to 0$. In other words, if ϵ is small, the seller the seller produces high quality with high probability and the customer also receives high quality with high probability. Hence we have constructed an equilibrium of the repeated game, with an outcome arbitrarily close to the efficient outcome, provided that the noise is sufficiently small. (In the appendix we complete the proof by showing that the continuation strategies at unreached information sets are also optimal).

Proposition 5 *There exist sequential equilibria of the infinitely repeated game with outcomes which are arbitrarily close to efficiency provided that* $\epsilon \to 0$ *and* $\delta > \frac{\Delta C}{(P-C_L)(1-\epsilon)}$.

The construction used here is related to that employed by Sekiguchi (1997), who analyzes the infinitely repeated prisoners' dilemma with private signals. He constructs an equilibrium where each player randomizes, choosing in period one between two repeated game pure strategies — a cooperative *grim trigger* strategy and a strategy of playing *always defect*. For a certain class of prisoners' dilemma payoffs, Sekiguchi shows that the player of the *grim trigger* strategy will find it optimal to continue cooperating if she receives good signals, but will switch to defecting on receiving a bad signal. Hence one can construct a mixed strategy equilibrium with partial cooperation. However, as the discount rate δ is increased, *always defect* yields low payoffs against a *grim trigger* strategy. Hence a player will only be willing to choose *always defect* if the other player plays *always defect* with high probability. Such an equilibrium cannot be efficient.[15] Sekiguchi's solution is to divide the overall game into N separate repeated games, where game k is played in periods $k, N+k, 2N+k$, etc. This reduces the effective discount factor and the efficient payoff can be approximated arbitrarily closely, provided that the noise $\epsilon \to 0$ and $\delta \to 1$.

Clearly, dividing up the game and public randomization play similar roles, of reducing the effective discount factor. The difference is that with one-sided moral hazard, it is difficult to construct an equilibrium with *any* degree of cooperation the absence of public randomization, since it is very difficult to make the seller indifferent between producing high and low quality. In the absence of a public signal, such indifference is possible with infinite trigger strategies only if the equality (17) holds when $m = 1$. Nor does dividing up the game work since this would require that (17) hold with $m = 1$ when we replace δ by δ^N. Analysis of equilibria in the absence of public randomization, appears to be very complex — we are not able to construct mixed strategy equilibria with any degree of cooperation, and nor can we rule out such equilibria.[16] In contrast, with public randomization, construction of cooperative equilibria has proved remarkably simple.

7 Concluding Comments

We now offer a brief summary of some of the related literature. Note that the work of Compte (1994, 1996), Kandori & Matsushima (1994), Kandori (1991), Mailath and Morris (1997), and Sekiguchi (1997) has also been discussed in previous sections.

The difficulty in supporting efficient outcomes in repeated games with private signals

[15] As $\delta \to 1$, the payoffs in such an equilibrium must approach the minimax payoffs. This result holds more generally – with independent private signals, in any equilibrium where players adopt any symmetric grim trigger strategies (i.e. where they do not cooperate if they have ever deviated), as Compte (1996) shows.

[16] This difficulty in analyzing mixed strategy equilibria has also been noted by Mailath and Samuelson (1997) in the context of a similar model. Consequently they focus attention on the case where signals are public.

was first pointed out by Matsushima (1991), who considered pure strategy equilibria in infinitely repeated games with independent private signals. With pure strategies, the signals are uninformative, and each player's beliefs about his opponents' future actions do not change with the realization of the signal. Matsushima assumed that players adopt strategies where they do not vary their actions in response to signals unless they have a strict incentive to do so, and proved an anti-folk theorem — all pure strategy Nash equilibria satisfying this property require players to play a Nash equilibrium of the stage game in each period. Mailath and Morris (1997) consider a "convention" game, where each of two players must choose an action in $[0,1]$. Although the convention game has a continuum of equilibria, Morris and Mailath show that with finitely many repetitions, players must play Nash equilibria of the stage game in every period, if signals are private.

An example of equilibrium with some cooperation, in the context of a twice-repeated game, was provided by Kandori (1991), and subsequently, in earlier versions of this paper. As we have discussed in section 3, the two constructions embody quite different ideas. The first example of purely non-cooperative equilibrium[17] in the context of an infinitely repeated game with private signals is due to Sekiguchi (1997). Sekiguchi's work and the current paper are complementary, and both employ the idea of using mixed strategies to allow players to learn from their private signals. Our focus is somewhat different — we study the effects of the level of noise (ϵ) upon equilibrium outcomes in a given repeated game (finite or infinite) with a *fixed* discount factor. In our contexts — the twice-repeated game with two-sided moral hazard and infinitely repeated one-sided moral hazard — Sekiguchi's device of dividing the game does not work. Public randomization turns out to be a flexible tool for supporting efficient outcomes. It also permits stronger results — our efficiency result does not require $\delta \to 1$, only that $\epsilon \to 0$ and that δ is greater than a critical value. By allowing for public randomization, it seems that Sekiguchi's result can be similarly strengthened.

Our results are related to and may have implications for the work on games with imperfectly observable commitment, introduced by Bagwell (1995). Bagwell observed that

[17] I.e. of an equilibrium without communication, as in Compte (1994) and Kandori-Matsushima (1994).

the slightest amount of imperfect observation destroyed a Stackelberg leader's advantage from pre-commitment, since the Stackelberg equilibrium was no longer a pure strategy equilibrium. van Damme and Hurkens (1995) showed that with one leader and one follower, the Stackelberg equilibrium could always be approximated by an equilibrium in mixed strategies, thereby ensuring that the sequential equilibrium correspondence was lower-hemicontinuous. Guth, Kirchsteiger and Ritzberger (1996) show, via an example, that lower-hemicontinuity is not ensured with more than one follower.

We conclude with an analogy which may be instructive. Dynamic games where players have *private information* about past events are yet to be fully understood. At first sight, these games bear a strong resemblance to Rubinstein's (1989) electronic mail game, and related infection arguments. Rubinstein's example shows that if the messages are noisy, exogenous and privately observed, players will not be able to condition their behavior on these messages.[18] The repeated games we discuss differ in one respect — signals are no longer exogenous, since players may influence them via their actions. Players may adopt two devices — *individual randomization*, so that each player is uncertain about his opponent's pure strategy, and *collective randomization*, which provides a favourable environment for individual randomization.These devices suffice to ensure very different results from Rubinstein's. While we are as yet far from a general theory of these games, we hope that the ideas suggested here may have a role in the development of a such a theory.

8 Appendix

Proof of Proposition 3:

Assume that $(\sigma_1(\epsilon),\, \sigma_2(\epsilon))$ is a mixed Nash equilibrium of $G^2(\epsilon, \rho)$ that is approximately efficient, i.e. the total payoff is approximately $\tilde{V}_H - C_H + V_L - C_L$.

[18]Infection arguments apply in other static contexts with imperfect information. For example, Carlsson and van Damme (1993) show that a small amount of payoff uncertainty selects a unique equilibrium in games with a multiple equilibria. See also Morris, Rob and Shin (1995). The analogy between repeated games with private signals and infection arguments has been explored further by Mailath and Morris (1997).

Define first the set Θ of *good* pure strategies in the repeated game, where a good strategy plays C in period 1, and responds to the signal c by playing D in period 2. $\Theta = \{(f_i, g_i) : f_i = C \text{ and } g_i(c) = D\}$. If the outcome of any mixed strategy is to be approximately efficient, then both players must be playing good strategies with probability close to one.

If player i is playing a good strategy with probability close to one, then any strategy of player j which plays E in period 1 is strictly inferior. Hence neither player plays E with positive probability in period 1.

If player i is playing C or D in period one, the first period payoff gain to player j from playing D rather than C in period 1 equals ΔC. Since player j plays C in period one with positive probability in equilibrium, this must not be an inferior choice. Hence if j plays D, he must suffer a loss in period 2 of at least ΔC. Hence player i must be playing a strategy which rewards the signal c and punishes the signal d. Call any such strategy α — i must assign positive probability to a pure strategy α. Since this argument applies for $i = 1, 2$, α is in the support of both players' strategies.

Let α' be a pure (good) strategy which plays C in period one, and responds to signal d by playing D in period two, i.e. this strategy does not punish after d.

Define the set Ξ of *bad* strategies as follows – any strategy from Ξ plays D in period one, and responds to the signal c by playing E. We now show that if player i assigns positive probability to α, then player j must assign positive probability to a bad strategy. We do this by showing that if no bad strategy is in the support of player $j's$ mixed strategy, then α is strictly inferior to α'.

Assume that no bad strategy is in the support of player $j's$ mixed strategy. Note that against $\sigma_j(\epsilon), \alpha$ and α' yield the same expected payoff in the first period, and also in the second period when i receives signal c. Hence, condition on j playing $\sigma_j(\epsilon)$, i playing α or α' and i receiving signal d. There are now two possibilities: player j is playing a pure strategy in the support of $\sigma_j(\epsilon)$ with $f_j = D$ or with $f_j = C$.

In the first case ($f_j = D$), since j is not playing a bad strategy and since he gets c with probability $(1 - \epsilon)$, he is most likely to play D. Consequently, in this case α' yields strictly more than α.

In the case when $f_j = C$, both players chose C in the first period and fig. 3 shows that j received signal c with probability $(1-\rho)(1-\epsilon)$. Since j is playing a good strategy with probability close to one, he continues with D with probability close to one after receiving signal c. We therefore conclude that, conditional on i receiving signal d and $f_j = C$, i believes that j will play D with probability approximately $1-\rho$ or more. Hence if $1-\rho > \mu^*$, then α' is strictly better than α in this case as well.

We conclude that if j does not play a bad strategy and if $\rho < 1-\mu^*$, then α yields strictly more than α when ϵ is sufficiently small. Since α is the support of $\sigma_i(\epsilon)$ for each player i, each player j must be playing a bad strategy with positive probability when $\rho < 1-\mu^*$.

Now, if j plays a good strategy and i plays a bad one, then $i's$ payoff is $\tilde{V}_H - C_L + F$, and hence against $\sigma_j(\epsilon)$ a bad strategy yields approximately $\tilde{V}_H - C_L + F$ in equilibrium. Since the payoff to all pure strategies in $\sigma_i(\epsilon)$ must be equal in any mixed Nash equilibrium, this implies that neither player's payoff can be greater than $\tilde{V}_H - C_L + F$. Since the efficient outcome has a strictly greater payoff , it cannot be approximated by any mixed Nash equilibrium of $G^2(\epsilon, \rho)$, no matter how small ϵ is. ■

Proof of Proposition 5

We show that each player's continuation strategy is optimal at unreached information sets. First, we complete the definition of strategies in the text by specifying actions at unreached information sets, as follows. (As before, actions in any period in a single game of stochastic length depend only on past events in that game).

i) If the customer has played N in any previous period, then the customer always chooses N and the seller always chooses D.

ii) If the seller has ever chosen D in any period, she always chooses D.

iii) If the customer has always chosen B, her action in the current period (t) only depends upon her signal in the last period $(t-1)$. She plays B and continues in the cooperative phase if this signal was c. She plays N and continues in the punishment phase if this was d.

From (i) it is clear that both players are choosing optimally after any history where the customer has chosen N at any point, since this is common knowledge.

Consider histories where the customer has always chosen B. If the signal in $t-1$ was c, the seller must have played C in all previous periods. Hence it is optimal for the customer to play B and continue in the cooperative phase. If thesignal received in $t-1$ was d, the customer assigns probability no more than $\theta(\pi)$ to the seller having chosen C in $t-1$. From inequality (21), it is optimal to play D. Thus the behavior prescribed in (iii) is optimal.

Finally, (ii) is optimal since D is always an optimal action regardless of whether the customer is in the cooperative phase or the punishment phase. ■

References

[1] Abreu, D., D. Pearce and E. Stachetti, 1990, Towards a Theory of Discounted Repeated Games with Imperfect Monitoring, *Econometrica,*58, 1041-1064.

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

[3] Bhaskar, V., 1994, Repeated Games with Almost Perfect Monitoring by Privately Observed Signals, Discussion paper 11, Centre for Development Economics, Delhi School of Economics.

[4] Bhaskar, V. 1996, On History-Dependent Mixed Strategies: A "Counter-Example" to Harsanyi's Theorem, mimeo, University of St. Andrews.

[5] Bhaskar, V., 1998, Informational Constraints and Overlapping Generations: An Anti-Folk Theorem, *Review of Economic Studies,* 65, 135-149.

[6] Carlsson, H., and E. van Damme, 1993, Global Games and Equilibrium Selection, *Econometrica,* 61,989-1018.

[7] Compte, O., 1994, Communication in Repeated Games with Imperfect Private Monitoring, *Econometrica,* forthcoming.

[8] Compte, O, 1996, On Failing to Cooperate when Monitoring is Private, mimeo, CERAS-ENPC, Paris.

[9] van Damme, E., and S. Hurkens, 1994, Games with Imperfectly Observable Commitment, *Games and Economic Behavior,* forthcoming.

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

[11] Green, E., and R. Porter, Non-Cooperative Collusion under Imperfect Price Information, *Econometrica* 52, 87-100.

[12] Güth, W., G. Kirchsteiger and K. Ritzberger, 1996, Imperfectly Observable Commitments in n-Player Games, *Games and Economic Behavior,* forthcoming.

[13] Harris, C., P. Reny and A. Robson, The Existence of Subgame-Perfect Equilibrium in Continuous Games with Almost Perfect Information: A Case for Public Randomization, *Econometrica*, 63, 507-544.

[14] Harsanyi, J., 1973, Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points, *International Journal of Game Theory* 2, 1-23.

[15] Kandori, M., 1992, The Use of Information in Repeated Games with Imperfect Monitoring, *Review of Economic Studies*, 59, 581-593.

[16] Kandori, M., 1991, Cooperation in Finitely Repeated Games with Imperfect Private Information, mimeo, University of Tokyo.

[17] Kandori, M and H. Matsushima, 1994, Private Observation, Communication and Collusion, *Econometrica,* forthcoming.

[18] Mailath, G., and S. Morris, 1997, Repeated Games with Imperfect Private Monitoring: Notes on a Coordination Persepective, mimeo, University of Pennsylvania.

[19] Mailath, G., and L. Samuelson, 1997, Your Reputation is Who You're Not, Not Who You Are, mimeo, University of Pennsylvania and University of Wisconsin, Madison.

[20] Matsushima, H., 1991, On the Theory of Repeated Games with Private Information, Part I, *Economics Letters* 35, 253-256.

[21] Morris, S., R. Rob and H. Shin, 1995, p-Dominance and Belief Potential, *Econometrica,*63, 145-167.

[22] Myerson, R., 1986, Multi-Stage Games with Communication, *Econometrica,* 54, 323-358.

[23] Radner, R., R. Myerson and E. Maskin , 1986, An Example of a Repeated Partnership Game with Discounting and Uniformly Inefficient Equilibria, *Review of Economic Studies*, 53, 59-70.

[24] Rubinstein, A., 1989, The Electronic Mail Game: Strategic Behavior under "Almost Common Knowledge", *American Economic Review*, 79, 385-391.

[25] Sekiguchi, T., 1997, Efficiency in the Prisoners' Dilemma with Private Monitoring, *Journal of Economic Theory,*76, 345-361.

[26] Shapiro, C., and J. Stiglitz, 1984, Equilibrium Unemployment as a Worker Discipline Device, *American Economic Review,* 74, 433-444.

[27] Stigler, G., 1964, A Theory of Oligopoly, *Journal of Political Economy* 72, 44-61.