Correlated Equilibrium and Trigger Strategies with Private Monitoring

Jeffrey C. Ely *
Department of Economics
Northwestern University
2003 Sheridan Road
Evanston IL 60208
ely@nwu.edu

Version: February 17, 2002

In repeated games, simple strategies such as Grim Trigger, while strict equilibria when monitoring is perfect, can fail to be even approximate Nash equilibria when monitoring is private, yet arbitrarily close to perfect. That is, they fail to be robust to private monitoring. In this paper, it is shown that for a class of repeated Prisoners' Dilemma games these strategies, when viewed as (degenerate) correlated equilibria *are* robust. In particular, even when monitoring is private and conditionally independent, as the signaling noise goes to zero, there is a sequence of correlated equilibria converging to the Grim Trigger strategies. The correlation device uses an information structure akin to that of the e-mail game of Rubinstein (1989).

Key Words: correlated equilibrium, repeated games, private monitoring, email game.

1. INTRODUCTION

In repeated games with perfect monitoring, the history of play is assumed to always be common knowledge among the players. This structure allows a wide range of payoff vectors to be supported as equilibria using strategies of a simple form: "cooperative behavior" enforced by the threat of mutual punishment. A simple example of such a strategy is the "grim trigger" strategy in the repeated prisoners' dilemma. The threat of perpetual defection gives players a strict incentive to cooperate.

When players monitor one another by imperfect and privately observed signals, these simple strategies often fail to be equilibria. In fact, I show below (Theorem 1) that the only strict equilibria that remain approximate

*I thank Juuso Välimäki for conversations leading to this paper. Jennie Lee provided valuable research assistance. Two embarrassingly careful and mercifully anonymous referees provided numerous suggestions which greatly improved both the content and exposition. Financial support from NSF grant #9985462 is gratefully acknowledged.

Nash equilibria under any small perturbation of the monitoring structure are history-independent alternations among stage game Nash profiles.

Mixed strategies can be used to support efficient cooperation for near-perfect monitoring structures. This has been shown in a number of papers, including Sekiguchi (1997), Bhaskar and Obara (2002), Piccione (2002), and Ely and Välimäki (2002). These equilibria, like all equilibria in mixed strategies, can be criticized as each player is expected to condition his play on information which is irrelevant for determining continuation payoffs. Bhaskar (1998) and Matsushima (1991) show that all strategies which ignore such payoff irrelevant details must yield static Nash play in every period.

These two sets of results leave a narrow space for equilibria that are robust to small monitoring imperfections. We must either be willing to accept mixed strategies, or accept strategies that are not even approximate equilibria for arbitrarily small perturbations of the model. In this paper I suggest an alternative justification for strict repeated game equilibria such as the grim trigger strategies. As (degenerate) *correlated* equilibria, they can be robust. That is, for small perturbations to the monitoring structure there may be correlated equilibria that yield approximately the same distribution over histories and hence the same payoffs.

In the next section I illustrate the idea in a simple 2-stage example. This is a version of the example first presented by Bhaskar and van Damme (2002) to shed light on the difficulties arising from private monitoring. The best subgame perfect equilibrium, which has a trigger-strategy structure, cannot be approximated by any Nash equilibrium that is robust to private monitoring. By contrast, I show that for any small monitoring perturbation there is a strict correlated equilibrium of the resulting game in which the original trigger strategy profile is played with probability close to 1. Thus, the trigger strategy equilibrium is a robust correlated equilibrium. In section 3, I provide a general negative result for strict equilibria of repeated games: they are never robust as *Nash* equilibria.

These results raise the following natural question: is every strict equilibrium robust as a correlated equilibrium? In section 4.1 I provide an example demonstrating that the answer is negative. A natural-looking trigger-strategy equilibrium of a two-stage game cannot be approximated by *any* equilibrium, correlated or not, when the monitoring technology is slightly perturbed. I point out at the end of this section the reason for the failure. The equilibrium in question implicitly requires one player to believe after seeing a deviation that his opponent is using a strategy that is not rationalizable. No such beliefs are possible when monitoring is imperfect.

Section 5, the heart of the paper, investigates the possibility that simple trigger strategies are robust correlated equilibria of repeated games. Specifically, I study a class of repeated prisoners' dilemma games and the standard grim-trigger equilibrium. Section 5.1 presents a family of strict correlated equilibria in which unilateral deviation occurs with positive probability in every period, but which nevertheless approximate arbitrarily well the path of play of the grim trigger profile. In Section 5.2 I show that these equilibria are robust to private monitoring. Finally, in section 5.3 I show how to embed these into a larger correlated equilibrium in which the actual grim-trigger profile is played with probability arbitrarily close to 1.

The prisoners' dilemma games studied in section 5 satisfy a certain restriction on the payoffs. This restriction is necessary for the type of correlated equilibria I employ. In section 6 I discuss the role that the restriction plays. While it appears that without this restriction, there are no correlated equilibria with a similar structure, it remains an open question whether grim trigger is robust in these games.

2. TWO-STAGE EXAMPLE

Versions of the following two-stage example where $M-1>g$ have been useful in illustrating the strategic features of repeated games with private monitoring, see especially Bhaskar and van Damme (2002).

	C	D
C	$1,1$	$-l,1+g$
D	$1+g,-l$	$0,0$

stage 1

	C	D
C	M,M	$0,0$
D	$0,0$	$1,1$

stage 2

FIG. 1 Stage 1 is a prisoners' dilemma: $g,l>0$. Stage 2 is a coordination game.

In the first stage, the players' play a prisoners' dilemma game. The coordination game played in the second stage is intended to capture the possibility of using multiple continuation equilibria in a repeated game to enforce cooperation in the first stage.

In each stage, each player chooses an action from the set $\{C,D\}$. Payoffs are the sum of the utilities in the two stages. After the first stage, each player observes a noisy signal of the action chosen by the other. The chosen action and the realized signal comprise the player's private history at the second stage. The set of signals is the same as the set of actions and each player is assumed to observe a correct signal with probability $(1-\varepsilon)$ and an incorrect signal with small probability $\varepsilon>0$. To avoid confusion, signals will be denoted by lower-case letters, i.e. c or d. To keep the example simple, we will assume independent monitoring: the probability of an incorrect signal is independent across players.

A *pure strategy* s_i for player i is a choice of action in the first stage and a choice of action in the second stage for each possible private history. For any strategy profile s, let $V_i(s)$ be player i's expected payoff. A *correlated strategy profile* $\tau=(\Omega_1,\Omega_2,\mu)$ consists of a set of types Ω_i for each player, a joint probability distribution μ over type profiles Ω and a pure strategy $s_i(\omega_i)$ for each type ω_i of each player i. A correlated profile is a *correlated equilibrium* if for each player i and each type ω_i such that $\mu(\omega_i) > 0$,

$$\sum_{\omega_j \in \Omega_j} \mu(\omega_j|\omega_i) V_i(s_i(\omega_i), s_j(\omega_j)) \geq \sum_{\omega_j \in \Omega_j} \mu(\omega_j|\omega_i) V_i(\hat{s}_i, s_j(\omega_j))$$

for each alternative pure strategy $\hat{s}_i$ for player i. If the measure μ assigns probability one to some type profile, then the equilibrium is a *pure-strategy equilibrium.* If μ is a non-degenerate distribution, but the two players' types are independent under μ, then the equilibrium is a *mixed-strategy Nash equilibrium.*[2]

There is no pure strategy equilibrium in which (C, C) is the outcome in the first stage. A player has an incentive to play C in stage one only if he expects that a choice of D will be punished by a lower payoff in the second stage. Such a punishment can be implemented only by a strategy that plays C in the second stage if and only if the signal c is observed. Therefore, if (C, C) is to be part of a pure-strategy equilibrium, each player must expect his opponent to be playing such a strategy. But with independent monitoring, this cannot be an equilibrium: after playing C, player 1 knows that player 2 has almost certainly seen a signal of c and will therefore almost certainly play C in the second stage. Against such beliefs, 1's best reply is to play C in the second stage. (It is optimal to play C in the second stage against any beliefs assigning probability at least $1/M + 1$ to the opponent playing C.) And player 1 has these beliefs regardless of the signal he observes: since monitoring is independent, 1's signal gives no further information about the signal observed by 2.

Let g be the strategy that plays C in stage one and in the second stage plays C in response to the signal c and D in response to the signal d. The above argument implies that the strategy profile (g, g), which is a strict subgame-perfect equilibrium when monitoring is perfect, fails to be an equilibrium when monitoring is imperfect, private and independent. In fact, there is no Nash equilibrium close to (g, g) and the set of Nash equilibrium (pure or mixed) payoffs is bounded away from efficiency, even as ε, the probability of an incorrect signal, goes to zero. This has been proven by Bhaskar and van Damme (2002). In this sense, the strict equilibrium (g, g) is not robust to private monitoring.

Consider the strategy b which plays D in each stage regardless of history. I will describe a correlated strategy profile in which each player plays g with high probability, but believes that the other is playing the strategy b with positive probability. With these prior beliefs, a signal of d after stage one,

[2]I have not imposed any sequential rationality conditions. It is well known that in games with full-support monitoring, Nash equilibrium is equivalent in all relevant respects to sequential equilibrium, see Sekiguchi (1997) for example. A similar equivalence applies to correlated equilibrium.

while uninformative about the signal observed by the other, nevertheless conveys some relevant information: that the opponent most likely played D in the first stage and therefore will continue with his strategy b and play D in the second stage. In response to such beliefs, it will be optimal to play D in the second stage. Thus, each player will condition his play on his private signal in a way that provides the appropriate first-period incentives.

To make this an equilibrium, the players must be willing to play b when called upon to do so. Correlation makes this possible without creating the inefficiencies inherent in mixed-strategy Nash equilibrium. In fact, in the correlated equilibrium I present below, it is strictly optimal to play C when called upon to do so, and D is played with vanishing probability in the limit as ε goes to zero. This implies that the strategy profile converges to the pure Nash profile (g, g) and that the equilibrium payoffs converge to the efficient point. In this sense, the strict equilibrium (g, g) which fails to be robust to private monitoring within the class of Nash equilibria, *is* robust as a (degenerate) correlated equilibrium.

The information structure used in the correlated equilibrium is as follows. Each player has two types, g and b. The probability distribution over type profiles is illustrated in the following matrix where $x = \frac{M\varepsilon}{(1-\varepsilon)}$

	g	b
g	1	x
b	x	$x^{1/2}$

FIG. 2 The correlation structure

The entries are relative probabilities. The probabilities are obtained by dividing each value by the sum of all values. We will analyze the correlated strategy profile in which each player plays the pure strategy corresponding to his realized type. As mentioned above, as ε approaches zero, the probability assigned to the profile (g, g) by this correlated profile approaches 1.

The crucial feature is that each player, regardless of his type, will respond to a d signal by playing D in the second stage. For a player of type b this is immediate: such a type plays D in stage 1 and therefore knows that with high probability his opponent will observe a d signal. Since both strategies respond to d with D, type b will be nearly certain that the opponent will be playing D in the second stage. In fact, because of the assumption of independent monitoring, this argument shows that a player of type b will play D in the second stage independent of his first stage signal.

For a player of type g, a d signal will be strong evidence that his opponent was of type b and will therefore play D in the second stage. To verify this, recall that D is a second stage best-reply to any belief that the opponent will play D with probability at least $\frac{M}{1+M}$. The probability that the opponent will play D in stage 2 is at least the probability that the opponent is of type b. By Bayes' rule, after observing the signal d in the first stage, the conditional probability that the opponent is of type b is

$$\frac{x(1-\varepsilon)}{\varepsilon + x(1-\varepsilon)} = \frac{M\varepsilon}{\varepsilon(1+M)} = \frac{M}{1+M}$$

Thus a type g player will also respond to signal d by playing D in the second stage.

Next, a player of type g who plays C in stage one and observes the signal c, should play C in the second stage. For this to be optimal, the conditional beliefs about the opponent's second stage action must assign at least probability $\frac{1}{1+M}$ to C. In equilibrium the opponent plays C when the opponent is of type g and saw a signal c. The conditional probability of this event for a type g who also sees c is

$$\frac{(1-\varepsilon)^2}{(1-\varepsilon) + x\varepsilon}$$

which is arbitrarily close to 1 and hence greater than $\frac{1}{1+M}$ for ε sufficiently small.

We have shown that second stage behavior prescribed by the strategy is optimal. Thus, each player can expect his opponent to play D in response to a d signal, and C in response to a g signal whenever his opponent is of type g. To complete the argument that the profile is a correlated equilibrium, we must show that first stage behavior is optimal. The straightforward calculations are omitted, the intuition is simple: when ε is sufficiently small, each player is nearly certain that his opponent is of the same type as he (the ratio of the diagonal to the off-diagonal elements of figure 2 approaches 1 as ε approaches zero.) Also each player is nearly certain that his opponent will observe a correct signal. Thus, for example, a player of type g believes his opponent is also g and will play C in stage 1 and with high probability punish a choice of D in stage 1. He is therefore also willing to play C.

To summarize the analysis in this section, punishments can be enforced only if players interpret the signal d as evidence that their opponent will play D in the second stage. This is impossible in a pure-strategy equilibrium. On the other hand, randomization allows players to believe that b is actually played with positive probability and therefore respond to d by punishing.

Inducing randomization in *Nash* equilibrium imposes incentive constraints that prevent even approximate efficiency: each player must be indifferent between b and g. Suppose payoffs are close to the efficient level $M+1$. Then it must be that the strategy g is played with probability close to one. But this means that b will almost surely be punished and will earn at most $g+2 < M+1$. Thus b does worse than g contradicting the requirement of mixed-strategy Nash equilibrium. By contrast, for small error probabilities there are correlated equilibria arbitrarily close to the efficient pure-strategy equilibrium. This is because correlation obviates these indifference constraints. Indeed, in the correlated equilibrium presented above, each type's strategy is strictly optimal.

Correlation in the strategy profile plays a different role than other forms of correlation used to obtain efficiency in games of this sort. Both Bhaskar and van Damme (2002) and Mailath and Morris (2002) consider correlation in the monitoring technology. This allows players to coordinate their second-stage play using their correlated histories. With sufficient correlation, there is a nearly efficient pure-strategy equilibrium. By contrast, correlation in the strategies can generate the necessary belief revision to enforce punishments even when monitoring is independent.

Correlation in the form of sunspots was also considered by Bhaskar and van Damme (2002). Sunspots allow players to agree to "forgive" any first-stage deviation with an appropriately chosen probability. This lowers the punishment to playing b, relaxing the incentive constraints which caused mixed-strategy Nash equilibria to be inefficient. As evidence of the distinct roles played by sunspots and the correlation device used above, note that it is crucial for the Bhaskar and van Damme (2002) equilibrium that the sunspots not be observed until stage 2 actually arrives. On the other hand, for the correlated equilibrium constructed above, it is necessary that the outcome of the correlation device be observed before play begins.

3. REPEATED GAMES AND STRICT EQUILIBRIA

The stage game is a finite strategic form game with action set A_i, and mixed actions ΔA_i for each player i. After each period of play, each player i observes a private signal from the set Σ_i. Assume that the set of signals for i coincides with the set of action profiles A_{-i} among the players other than i. Conditional on each action profile $a \in \prod_i A_i$, a signal profile σ is drawn from the set $\prod_i \Sigma_i$ according to the distribution $m(\cdot|a)$. Perfect monitoring can be represented by the family of signal distributions $m(\cdot|a)$, such that for each a, $m(\sigma^a|a) = 1$, where $\sigma_i^a = a_{-i}$.

A history for player i of length t is an element of $A_i^t \times \Sigma_i^t$. Let H_i^t denote the set of all t-length histories with $H_i^0 = \{\emptyset\}$. Finally H_i is the set of all histories for i. Strategies, denoted $s_i = \{s_i^t\}_{t \geq 1}$ are sequences of maps $s_i^t : H_i^{t-1} \to \Delta A_i$. A history h_i is on the (equilibrium) path under profile s if h_i arises with positive probability under s. A history h_i is *consistent* with strategy s_i if there is some profile s_{-i} such that h_i is on the path for (s_i, s_{-i}).

Continuation strategies after histories h_i^t are denoted $s_i(h_i^t)$. For any monitoring distribution, a strategy profile induces a well-defined discounted payoff vector. Denote by $V_i(s)$ the discounted payoff to player i when the strategy profile is s. If $(h_1^t, \ldots, h_n^t)$ is the profile of histories observed through period t, then $V_i(s_1(h_1^t), \ldots, s_n(h_n^t))$ is player i's continuation payoff. Conditional on having observed history h_i^t under profile s, player i may not be able to infer the histories observed by the other players, and hence their continuation strategies. However, when h_i^t is on the path, i has a well-defined belief over these histories. For such a history h_i^t, denote by $V_i(\tilde{s}_i|s, h_i^t)$ the expectation of $V_i(s_1(h_1^t), \ldots, \tilde{s}_i, \ldots, s_n(h_n^t))$ with respect to the belief over opponents' histories derived from the profile s, conditional on i's observed history h_i^t.

For the case of perfect monitoring, we will consider *strict* (Nash) equilibria.

DEFINITION 1. A strategy profile s is a *strict* equilibrium of the game with perfect monitoring if for each i and each history $h_i \in H_i$ on the path, $V_i(s_i|s, h_i) > V_i(\tilde{s}_i|s, h_i)$, for any continuation strategy $\tilde{s}_i$ whose initial action differs from $s_i(h_i)$.

A strict equilibrium is one in which each player has a strictly optimal action after every history on the path of play. Note that the strategies used by Fudenberg and Maskin (1986) to prove the folk theorem are strict equilibria. Let us now consider perturbations of the perfect monitoring structure. Say that a monitoring structure m is ε-perfect if it has full support, (if $\forall\ \sigma, \forall\ a,\ m(\sigma|a) > 0$) and if for each $a \in A$, $m(\sigma^a|a) > 1 - \varepsilon$. I will show that apart from repetitions of static Nash profiles, any strict equilibrium under perfect monitoring fails to be an equilibrium even for monitoring structures arbitrarily close in this sense to perfect monitoring. In fact, they will fail to be even approximate equilibria in the following sense.

DEFINITION 2. Strategy s_i is a *δ-best-reply* to s_{-i} at history $h_i \in H_i$ if either h_i is off the path of play under $s = (s_i, s_{-i})$ or $V_i(s_i|s, h_i^t) > V_i(\tilde{s}_i|s, h_i^t) - \delta$ for any continuation strategy $\tilde{s}_i$ whose initial action differs from $s_i(h_i^t)$.

The above is a notion of approximate best-response defined at a particular history. I will say that an equilibrium under perfect monitoring is weakly robust to private monitoring perturbations if for every consistent history, there is a small enough perturbation such that each player's original strategy remains an approximate best-response at that history.

DEFINITION 3. An equilibrium s of a repeated game with perfect monitoring is *weakly robust* if for each i, $\delta > 0$ and h_i consistent with s_i, there is $\varepsilon > 0$ such that under any ε-perfect monitoring technology, s_i is a δ-best-reply to s_{-i} at h_i.

This is a weak form of robustness because, first, the strategies are required only to be approximate best-responses after the perturbation, and second, the size of the perturbation can depend on the history in question. In particular, weak robustness does not require the existence of a uniform bound on the perturbation size which guarantees that the strategies remain approximate best-responses at all histories simultaneously. Yet, strict equilibria fail even this weak form of robustness. The argument is related to the one made by Matsushima (1991).

THEOREM 1. *If s is a strict equilibrium under perfect monitoring and s is weakly robust, then s consists of history-independent plays of static Nash profiles.*

Proof. We will show that the hypotheses of the theorem imply that for each t, each strategy in the profile s plays the same action after every t length history consistent with the strategy. Note that this implies s consists of history-independent plays of stage-game Nash profiles.

Suppose, on the contrary that there is some player i and some stage t such that there are two consistent t length histories for which s_i prescribes different actions. Assume wlog that t is the first stage in which this is true for any player i.

Since s is a strict equilibrium, each player has a unique optimal action after every history on the path of play. This implies in particular that when monitoring is perfect, there is a deterministic path of play. Let $\bar{a}_j$ be the sequence of actions taken by player j in the first t periods in this deterministic path, and $\bar{a}_{-j}$ the sequence of action profiles among the players other than j.

When monitoring is ε-perfect, the probability is at least $(1-\varepsilon)^t$ after any given stage that the realized signal profile matches the chosen action profile. Therefore, the unconditional probability is at least $(1-\varepsilon)^t$ that the profile of histories observed by i's opponents is $h^t_{-i} = ((\bar{a}_j, \bar{a}_{-j}))_{j\neq i}$. Since i plays $\bar{a}_i$ in the first t periods with probability 1, this probability is unchanged when i conditions on his own history of actions. Choose a monitoring technology that is *independent*, i.e. that for each action profile, a and signal profile σ, $m(\sigma|a)$ is the product of the marginal distributions $m^j(\sigma_j|a)$. Then conditional on the action profile sequence $\bar{a}$, player i's history of signals is uninformative of the signals observed by the other players. Since $\bar{a}$ occurs with probability 1, we can conclude that conditional on every t length history on the path of play, player i assigns probability at least $(1-\varepsilon)^t$ to the history h^t_{-i} for his opponents.

Since i is conditioning his action choice on his history in stage $t+1$, there must be at least one history h^t_i consistent with s_i such that $s^{t+1}_i(h^t_i) \neq s^{t+1}_i(\bar{a}_i, \bar{a}_{-i})$. Assume the monitoring technology assigns positive probability to h^t_i conditional on $\bar{a}$. (Since h^t_i is consistent, it is sufficient that the monitoring distributions have full support). Since s is strict, there exists a $d > 0$ such that

$$V_i(s_i(\bar{a}_i, \bar{a}_{-i}), s_{-i}(h^t_{-i})) - V_i(s_i(h^t_i), s_{-i}(h^t_{-i})) > d$$

Choose δ to satisfy $0 < \delta < d$. By the argument above about conditional beliefs in ε-perfect monitoring structures, as ε goes to zero,

$$V_i^{\varepsilon}(\tilde{s}_i|s, \tilde{h}_i^t) \to V_i(\tilde{s}_i, s_{-i}(h_{-i}^t))$$

for any t-length history $\tilde{h}_i^t$, and continuation strategy $\tilde{s}_i$, where V_i^{ε} is the conditional expected payoff function for some ε-perfect monitoring structure. Therefore, we can take ε small enough so that

$$V_i^{\varepsilon}(s_i(\bar{a}_i, \bar{a}_{-i})|s, h_i^t) - V_i(s_i(h_i^t)|s, h_i^t) > \delta$$

implying that s_i is not an δ-approximate best-response at h_i^t.

We conclude that no player conditions his action choice on any consistent history and this concludes the proof. QED

This result is complementary to the positive results of Mailath and Morris (2002). They prove a folk theorem using strict equilibria that are robust to sufficiently "public" private monitoring imperfections. Theorem 1 shows that strict equilibria cannot be robust to *all* private monitoring perturbations. In Ely and Välimäki (2002), it is shown that there are efficient mixed strategy equilibria of infinitely repeated games that are robust to all monitoring imperfections. These strategies, like all mixed strategies, can be criticized because they require players to condition their behavior on payoff-irrelevant details.[3] In section 5, I show that in some games, the robustness of the simple and strict grim trigger equilibria can be recovered if we view them as degenerate correlated equilibria.

4. CORRELATED EQUILIBRIA

In the previous subsection, it was shown that the only weakly robust strict equilibria of repeated games are the trivial ones. On the other hand, the example of section 2 shows that there are non-trivial strict Nash equilibria which can be approximated arbitrarily well by strict robust correlated equilibria. An infinitely repeated example will be presented in section 5 below. It is natural to ask whether strict Nash equilibria can *always* be so approximated. In the next subsection I provide an example demonstrating that the answer is negative. This example adds to those of Bagwell (1995) and Bhaskar and van Damme (2002). Bagwell (1995) first illustrated a strict pure-strategy Nash equilibrium which was not robust to monitoring imperfections. This equilibrium however can be approximated by robust mixed-strategy Nash equilibria. The example of Bhaskar and van Damme (2002) improves upon this because the strict pure-strategy Nash equilibrium they focus on cannot be approximated by any robust mixed-strategy Nash equilibrium. However, as was shown in section 2 this equilibrium can be approximated by robust strict correlated equilibria.

[3]The proof of theorem 1 makes it clear that they must.

4.1. A Strict Equilibrium That is Not Robust

The following is an example of a strict pure-strategy Nash equilibrium which *cannot* be approximated by any robust strict correlated (and hence Nash) equilibrium.[4]

The example adds a single action to the two stage game from section 2. We use uppercase for actions, and the corresponding lower case for the associated monitoring signals.

	C	D
C	$1,1$	$-4,4$
D	$4,-4$	$0,0$
W	$2,0$	$-7,7$

stage 1

	C	D
C	$6,6$	$0,0$
D	$0,0$	$1,1$

stage 2

FIG. 3 Playing (C, C) in stage 1 is not robust.

The best Nash[5] equilibrium of this game plays (C, C) in stage 1 supported by the promise of (C, C) in stage 2 and the threat of (D, D) in stage 2 if any deviation occurs. This is in fact a strict equilibrium in the sense of Definition 1. I claim that any strict correlated equilibrium in which (C, C) is played with probability close to 1 in stage 1 is not weakly robust.

Consider a strict correlated equilibrium under perfect monitoring in which the total probability of (C, C) in stage 1 is at least $1 - \epsilon$ for $\epsilon > 0$. By the law of total probability, there must be a positive probability type τ_1 of player 1 conditional on which the probability is at least $1 - \epsilon$ that (C, C) will be played in stage 1.

Type τ_1 is tempted to play W in stage 1. Therefore τ_1 must assign positive probability to a type τ_2 of player 2 who plays a strategy with the following features: "play C in stage 1 and respond to a signal of w by playing D in stage 2." To prove this, suppose the contrary: that every τ_2 which has positive conditional probability for τ_1 and which plays C in stage 1, responds to w by playing C. Then ϵ is an upper bound on the total conditional probability for τ_1 that action W will be punished by a response of D. This would mean that for ϵ small, a better reply for τ_1 would be to play W in stage 1 and C in stage 2, a contradiction.

Type τ_2 cannot assign positive conditional probability to a type τ_1' of player 1 which plays W in stage 1. We prove this in 2 steps. First, any strategy for player 1 which plays W in stage 1 and responds to a signal c with the action D in stage 2 is strictly dominated. Indeed if player 2

[4]Strictly speaking, weak robustness has been defined for Nash equilibria of repeated games, whereas we will now investigate weak robustness of correlated equilibria of a dynamic game. The definitions are readily translated and no confusion should result.

[5]In fact the best correlated equilibrium.

were playing C in stage 1, such a strategy could earn at most 3, whereas any strategy which plays D in stage 1 earns at least 4. And if player 2 were playing D in stage 1, such a strategy could earn at most -1, while any strategy which plays D in stage 1 earns at least 0. This means that, in our correlated equilibrium, any type of player 1 which plays W in stage 1 must respond to signal c with action C in stage 2. Thus if, τ_2 assigned positive conditional probability to such a type, τ_2 would expect action C in stage 2 after his own play of C and the signal w in stage 1. But this would be a contradiction since τ_2 responds to w with D.

Recall that the correlated equilibrium in question is strict. Let $2\delta > 0$ be the difference in expected payoffs for type τ_2 between actions D and C following the outcome (C, c) in stage 1. Let ϕ_C be the conditional beliefs of type τ_2 over the action to be played by player 1 in stage 2 after the outcomes (C, c). Note that this is well-defined since (C, c) is on the path of play for τ_2.

We will now construct the monitoring perturbation which exposes the non-robustness of the given correlated equilibrium. For arbitrary $\varepsilon \in (0,1)$, and $\lambda \in (0,1)$, consider an independent ε-perfect monitoring technology where

$$\frac{m^2(w|(C,C))}{m^2(w|(D,C))} = \lambda$$

Let $\tilde{\phi}_C$ and $\tilde{\phi}_D$ be the probability distribution over the action to be played by player 1 in stage 2 conditional on τ_2 and the chosen actions in stage 1 being (C,C) and (D,C) respectively. Note that $\tilde{\phi}_C$ approaches ϕ_C as ε approaches zero.[6] Let t be the probability type τ_2 assigns to the event that player 1 has actually played C in stage 1 conditional on τ_2 playing C and observing signal w. Note that t approaches 1 as λ approaches 1.

The conditional belief of type τ_2 over the action to be played by player 1 in stage 2 after playing C and observing signal w is

$$t\tilde{\phi}_C + (1-t)\tilde{\phi}_D$$

Because the expected payoff difference between actions C and D is a continuous function of this belief, we choose ε sufficiently small and λ sufficiently close to 1 to make this belief close enough to ϕ_C so that C earns in expectation at least δ more than D. Thus, type τ_2, which plays D after this history is not playing a δ-best reply. This shows that the correlated equilibrium is not weakly robust.

The argument above identifies the flaw in the cooperative equilibrium of this game. Player 2 is required to punish a play of W by playing D in stage 2. But player 2 can only rationally play D in stage 2 if he expects 1

[6]The only reason $\tilde{\phi}_C$ does not already equal ϕ_C is that player 1 is conditioning his second stage action on his signal, and when $\varepsilon > 0$ this signal can have errors. But as ε approaches zero, the signal is almost certainly equal to the action actually chosen by player 2.

to follow up his play of W with D. But the repeated game strategy which plays W and then D is strictly dominated. The cooperative equilibrium hinges on player 2 believing after stage 1 that player 1 was playing a non-rationalizable strategy! When the monitoring has full-support, Bayes rule determines player 2's belief. In particular, it constrains 2's belief to assign positive probability only to those strategies that are actually being played by player 1 in equilibrium. In particular, 2 must believe that 1 is playing a rationalizable strategy.

5. REPEATED PRISONERS' DILEMMA

In this section I construct an approximately efficient correlated equilibrium of the infinitely repeated prisoners' dilemma with private monitoring. Two complications arise in extending the correlated equilibrium construction exemplified in Section 2 to the infinitely repeated game. Recall that when monitoring is private and independent, in order for players to have a strict incentive to respond to bad signals, they must believe that their opponent is actually defecting with positive probability. Since equilibrium will require that players punish bad signals in any period such a signal is observed, it follows that correlated equilibrium of the type used in section 2 must have each player unilaterally defecting with positive probability in *every* period. I show that such a correlated equilibrium can be constructed provided the gain to cooperation is large enough relative to the gain to unilateral defection, and moreover that this equilibrium is strict when players are sufficiently patient.

The second complication arises when showing robustness of this type of equilibrium to private monitoring imperfections. Since unilateral deviation occurs with positive probability in every period, beliefs following a deviation from mutual cooperation are determined by Bayes' rule and are thus continuous in the monitoring noise. Thus beliefs at any information set can be made as close as necessary to the perfect-monitoring beliefs by choosing the noise sufficiently small. The difficulty is in showing that this can be done uniformly across the infinite set of histories, i.e. that there exists a sufficiently small monitoring perturbation so that beliefs after every information set are close enough to the original beliefs so that best-responses are unchanged. This is accomplished by using a particular type of "stationary" correlation structure.

Stage game payoffs are as in stage one of figure 1, with $g < 1$. The reason for this payoff restriction will be explained below. After each stage, each player privately observes a signal of the action chosen by the opponent. For simplicity I assume there are two possible signals $\Sigma = \{c, d\}$ corresponding to the two actions in the prisoners' dilemma. To simplify notation, I will focus on a particular *independent* private monitoring technology. With probability $1 - \varepsilon$, player i observes the signal corresponding to the action actually taken by player j; with probability ε, i observes an *error*. These probabilities are independent for the two players. The players' share a common discount factor δ and seek to maximize expected discounted payoffs. This parameterized family of repeated games will be denoted $G^{\infty}(\delta, \varepsilon)$.

5.1. Strict Correlated Equilibria

I will start by constructing a correlated equilibrium of the repeated game with *perfect* monitoring. This equilibrium will be strict and will have the property that unilateral defection occurs with positive probability in every period. This will imply that beliefs following a defection will be well-defined via Bayes rule. In equilibrium, a unilateral deviation by the opponent will signal that the opponent will continue defecting forever thereafter. Because these beliefs will be defined by Bayes' rule, they will change continuously with small changes in the monitoring technology. In fact, I will show that this continuity in beliefs is uniform across histories. Thus, since the original equilibrium is strict and beliefs uniformly continuous, the equilibrium will be robust to small private monitoring imperfections.

The construction takes a slightly different form in the two cases $l \geq 1$, $l < 1$. I will illustrate the $l \geq 1$ case in detail and briefly discuss the $l < 1$ case at the end. Throughout, maintain the assumptions $l \geq 1$, $g < 1$. Each player's type ω_i in the correlation device will be drawn from the set $\mathbb{N} = \{1, 2, \ldots\}$. Choose an integer k so that $k - 1 \leq l$ and $2k > 1 + g + l$. Note that this is always possible since $l \geq 1 > g$ (For example, choose k to be the smallest integer no smaller than l.) Fix $\beta < 1$ and define the following function on $\mathbb{N} \times \mathbb{N}$:

$$f(\omega_1, \omega_2) = \begin{cases} (1-\beta)^{\omega_1+\omega_2} & \text{if } \omega_2 \in \{\omega_1 - k, \omega_1, \omega_1 + k\} \\ z(1-\beta)^{\omega_1+\omega_2} & \text{otherwise} \end{cases}$$

Let μ be the probability measure on $\mathbb{N} \times \mathbb{N}$ defined by

$$\mu(\omega_1, \omega_2) = \frac{1}{M} f(\omega_1, \omega_2)$$

where

$$M = \sum_{\omega_1, \omega_2} f(\omega_1, \omega_2) < \infty$$

For $z = 0$, each type θ_i of player i considers possible only 3 types of the opponent , $\theta_i - k, \theta_i, \theta_i + k$. This information structure is partially illustrated in figure 4 for the case $k = 2$.[7] When z is positive but small, every type of the opponent is possible, but the "non-diagonal" types are relatively less likely.

[7]The figure depicts only the subset of the type space consisting of even numbered types. When $z = 0$, this set is in fact a common-knowledge cell of the information structure.

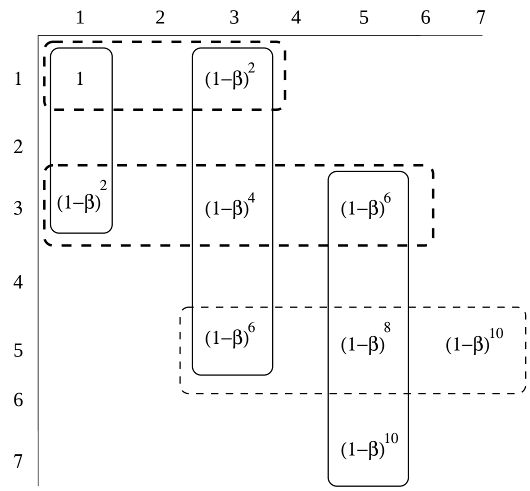

FIG. 4 Correlation device

The strategy *grim* is the repeated game strategy which plays D after every history in which either D or d appears at least once and plays C after every other history. The strategy *defect* is the strategy which plays D after every history. Let g_t denote the repeated game strategy which plays grim through period $t-1$ and switches to defect in period $t+1$ and thereafter. Note that g_1 is *defect*.

The correlated strategy profile will specify that type $\omega_i = t$ should play g_t, define $s_i(t) = g_t$ for each i and each $t \in \mathbb{N}$. Then $(\mathbb{N} \times \mathbb{N}, \mu, s)$ is a correlated strategy profile which I will call $\tau(\beta, z)$.

I will use a shorthand notation to refer to private histories. The notation $(t, (C, c)^s)$, for example is a history for i in which i was assigned type $\omega_i = t$ by the correlation device, and the outcome in each of the first s stages was (C, c), i.e. i played C and observed the signal c from j.

Say that a correlated equilibrium is *uniformly strict* if for each player i there is a $d > 0$ such that for each type ω_i, for every private history consistent with $s_i(\omega_i)$, the payoff resulting from the equilibrium continuation strategy exceeds by at least d the payoff to any alternative continuation strategy which deviates in the first stage following the history in question.

LEMMA 1. *For every δ sufficiently large and $\epsilon > 0$, there exist $\bar{\beta}$ and $\bar{z}$ sufficiently small such that for all $0 < \beta < \bar{\beta}$ and $0 < z < \bar{z}$, the correlated strategy profile $\tau(\beta, z)$ is a uniformly strict correlated equilibrium of $G^{\infty}(\delta, 0)$ in which each player receives a payoff greater than $1 - \epsilon$.*

Proof. First note that for every $\delta < 1$, the payoff to $\tau(\beta, z)$ is a continuous function of β and converges to 1 as β approaches zero. Thus, for any

z, δ and ϵ, β can be chosen sufficiently small so that the payoff to $\tau(\beta, z)$ exceeds $1-\epsilon$.

To show that $\tau(\beta, z)$ is a uniformly strict correlated equilibrium for β and z sufficiently small, we show that attention can be restricted to 5 histories, the others being redundant.

First, consider a history in which at least one D or d has appeared. After *any* such history, each player assigns probability 1 to the opponent playing the continuation strategy *defect*. The equilibrium prescribes also continuing with *defect* and earning a continuation payoff of 0. Any strategy which deviates and plays C can earn at most $-l$. Thus, the payoff difference is at least l after all histories in this class.

Now we consider histories in which (C, c) has occurred in each stage. If (C, c) has occurred through s stages, then the opponent's type cannot be $\omega_j \leq s$, otherwise he would have played D and since monitoring is perfect, d would have been observed. The stationary structure of the correlation device implies that the conditional beliefs over the opponent's type after the history $(t, (C, c)^{t-n})$ are independent of the type t. Since these beliefs determine beliefs over the opponent's continuation strategy and therefore continuation payoffs; to show uniform strictness, it suffices to choose an arbitrary t and show strictness uniformly for all histories $(t, (C, c)^{t-n})$ with $n \geq 0$.

To show uniform strictness we must find those histories in this class where the payoff difference between conforming to the equilibrium continuation and deviating is minimized and show that this payoff difference is strictly positive. We will show that when z is small enough, only four histories in this class are relevant, namely $(t, (C, c)^{t-k-2})$, $(t, (C, c)^{t-2})$, $(t, (C, c)^{t-k-1}$ and $(t, (C, c)^{t-1})$ all others having strictly larger payoff difference than at least one of these.

For small z, the payoff difference for a history of the form $(t, (C, c)^{t-k-n})$ for $n > 2$ is strictly larger than for a history of the form $(t, (C, c)^{t-k-n+1})$. In either case the continuation prescribed by the equilibrium is to play *grim* until period t. Let V be the equilibrium continuation payoff after $(t, (C, c)^{t-k-n+1})$. For small z, after either of these histories, the payoff to any strategy which deviates and plays D is approximately[8] $1+g$. The equilibrium continuation payoff after $(t, (C, c)^{t-k-n})$ is approximately $1 + \delta V$ which is greater than V when δ is close enough to 1.

It follows that among all histories of this form, it suffices to consider $(t, (C, c)^{t-k-2})$. Note that the same argument implies that among all the histories of the form $(t, (C, c)^{t-n})$, for $n < k$, it suffices to consider $n = 2$.

We now analyze these four histories in turn. First consider the histories $(t, (C, c)^{t-k-2})$ and $(t, (C, c)^{t-k-1})$. After these histories, for small z, the conditional probabilities of the opponent's types are approximately as

[8]The conditional probability that the opponent will play continue cooperating is arbitrarily close to 1 when z is small.

follows

$$\begin{aligned}
\mathrm{Prob}(\omega_j = \omega_i - k|h) &= \frac{1}{1+(1-\beta)^k+(1-\beta)^{2k}} \\
\mathrm{Prob}(\omega_j = \omega_i|h) &= \frac{(1-\beta)^k}{1+(1-\beta)^k+(1-\beta)^{2k}} \\
\mathrm{Prob}(\omega_j = \omega_i + k|h) &= \frac{(1-\beta)^{2k}}{1+(1-\beta)^k+(1-\beta)^{2k}}
\end{aligned}$$

Note that each converges to 1/3 as β approaches 0.

Consider the history $(t,(C,c)^{t-k-1})$. Playing D in period $t-k$ gives maximum continuation payoff

$$\left[1 - \frac{1}{1+(1-\beta)^k+(1-\beta)^{2k}}\right](1+g)$$

while instead conforming to the equilibrium recommendation yields (approximately for small z)

$$\frac{1}{1+(1-\beta)^k+(1-\beta)^{2k}}\left\{-l+(1-\beta)^k\sum_{n=0}^{k-1}\delta^n+(1-\beta)^{2k}\left[\delta^k(1+g)+\sum_{n=0}^{k-1}\delta^n\right]\right\}$$

(The three scenarios are: the player loses l because the opponent's type is $t-k$, the player earns the cooperative payoff for k additional periods, then zero forever because the opponent's type is t, and the player earns the cooperative payoff for k additional periods and then the pre-emption payoff $1+g$ because the opponent's type is $t+k$.)

For large δ and small β the latter is larger than the former if

$$1/3(2k+(1+g)-l) > 2/3(1+g)$$

which holds because $2k > 1+g+l$.

Next consider the history $(t,(C,c)^{t-k-2})$. Defecting after this history yields approximately $1+g$ (for z small).

This is strictly worse than playing C and then playing D in period $t-k$ period provided

$$1+g < 1+\delta\left[1-\frac{1}{1+(1-\beta)^k+(1-\beta)^{2k}}\right](1+g)$$

For large δ this inequality is approximately

$$\frac{1+g}{1+(1-\beta)^k+(1-\beta)^{2k}} < 1$$

which holds for small enough β since $1+g<2$. Since we previously showed that playing D in period $t-k$ was strictly worse than following the equilibrium, we can conclude that playing D in period $t-k-1$ is also strictly worse.

Next consider the history $(t, (C,c)^{t-2})$. For small z, conditional beliefs are approximately

$$\begin{aligned} \text{Prob}(\omega_j = t|h) &= \frac{1}{1+(1-\beta)^k} \\ \text{Prob}(\omega_j = t+k|h) &= \frac{(1-\beta)^k}{1+(1-\beta)^k} \end{aligned}$$

The equilibrium calls for the player to cooperate in stage $t-1$ and defect in stage t. This gives approximately

$$1+\delta\left\{\frac{(1-\beta)^k}{1+(1-\beta)^k}(1+g)\right\}$$

By instead defecting in stage $t-1$, the player gets at most $1+g$ which for small β and high δ is strictly worse provided

$$1+g < 1+\frac{1+g}{2}$$

which holds because $g<1$.

Finally, consider the history $(t, (C,c)^{t-1})$, after which the equilibrium prescribes defecting forever and earning approximate payoff

$$\frac{(1-\beta)^k}{1+(1-\beta)^k}(1+g)$$

For z small, the best deviation will be to play instead g_{t+k-1}, attempting to gain the cooperative payoff from $\omega_j = t+k$ and pre-empting by defecting in period $t+k-1$. This gives approximate payoff

$$\frac{1}{1+(1-\beta)^k}\left\{(-l)+(1-\beta)^k\left[\delta^{k-1}(1+g)+\sum_{n=0}^{k-2}\delta^n\right]\right\}$$

This expression is strictly increasing in δ. Hence, because $\delta<1$, for conforming to be superior, it is sufficient that

$$(1-\beta)^k(1+g) \geq (1-\beta)^k\left[(k-1)+(1+g)\right]-l$$

and for β small enough, this follows by the definition of k: $l \geq k-1$.

QED

5.2. Independent Private Monitoring

We now consider the game $G^\infty(\delta,\varepsilon)$ for small ε. The main result is that the conditional beliefs implied by the correlated strategy $\tau(\beta, z)$ are continuous in ε, uniformly across all histories consistent with the strategy. Because the equilibrium is uniformly strict for $\varepsilon = 0$, this implies that $\tau(\beta, z)$ is an equilibrium of $G^\infty(\delta, \varepsilon)$ for all sufficiently small ε. That is, the equilibrium is robust to private monitoring.

LEMMA 2. *Let $\tilde{\nu}_i^\varepsilon(\cdot|h)$ represent i's conditional beliefs over j's continuation strategy under $\tau(\beta, z)$ in the game $G^\infty(\delta, \varepsilon)$. Then $\tilde{\nu}_i^\varepsilon(\cdot|h) \to \tilde{\nu}_i^0(\cdot|h)$ uniformly for all histories h consistent with the strategy.*

Proof. Begin with some notation and a preliminary result. Fix t and a player i. Define μ_t to be the conditional distribution over ω_j, $j \neq i$ for type $\omega_i = t$ of player i. Let F_s be the event $\{\omega_j > s\}$, and I the event $\{\omega_j \in \{\omega_i - k, \omega_i, \omega_i + k\}$.

Let $s < t$. If $s \neq t - k$ then

$$\frac{\mu_t(s)}{\mu_t(F_s)} = \frac{z(1-\beta)^{t+s}}{\sum\limits_{n>t+s,\, n\notin I} z(1-\beta)^n + \sum\limits_{n>t+s,\, n\in I} (1-\beta)^n}$$

which can be bounded uniformly as follows

$$\frac{z\beta}{1-\beta} < \frac{\mu_t(s)}{\mu_t(F_s)} < \frac{\beta}{1-\beta}$$

and if $s = t - k$ then

$$\frac{\mu_t(s)}{\mu_t(F_s)} = \frac{(1-\beta)^{t+s}}{\sum\limits_{n>t+s,\, n\notin I} z(1-\beta)^n + \sum\limits_{n>t+s,\, n\in I} (1-\beta)^n}$$

which can be bounded uniformly as follows

$$\frac{\beta}{1-\beta} < \frac{\mu_t(s)}{\mu_t(F_s)} < \frac{\beta}{z(1-\beta)}$$

Thus, for all $s < t$

$$\frac{z\beta}{(1-\beta)} < \frac{\mu_t(s)}{\mu_t(F_s)} < \frac{\beta}{z(1-\beta)}$$

Consider any history of the form $h = (t, (C, c)^s)$ for $s < t$. There are three types of events which could have lead to this history. First, the opponent's type could be some $\omega_j > s$ and there were no errors through period s. The probability of this event is

$$\mu_t(F_s)(1-\varepsilon)^{2s}$$

Second, the opponent's type could be $\omega_j > s$, but the opponent saw an erroneous d in some period λ, began defecting, but player j erroneously saw c in all periods from $\lambda + 1$ to s. The probability of this event is

$$\mu_t(F_s)\sum_{\lambda=1}^{s}(1-\varepsilon)^{2\lambda-1}\varepsilon^{s-\lambda+1}$$

Finally, the opponent's type could be $\omega_j \leq s$. In this case the first error was observed in some period no later than ω_j. The probability of this event is less than[9]

$$\sum_{m=1}^{s} \mu_t(m) \sum_{\lambda=1}^{m} (1-\varepsilon)^{2(\lambda-1)} \varepsilon^{s-\lambda+1}$$

Thus, using Bayes' rule, conditional on such a history, player i believes his opponent is using continuation strategy g_n for $n \geq 1$ with probability

$$\begin{aligned}
\tilde{\nu}_i^{\varepsilon}(g_n|h) &> \frac{\mu_t(n+s)(1-\varepsilon)^{2s}}{\mu_t(F_s)\left[(1-\varepsilon)^{2s} + \sum\limits_{\lambda=1}^{s}(1-\varepsilon)^{2\lambda-1}\varepsilon^{s-\lambda+1}\right] + \sum\limits_{m=1}^{s}\mu_t(m)\sum\limits_{\lambda=1}^{m}(1-\varepsilon)^{2(\lambda-1)}\varepsilon^{s-\lambda+1}} \\
&= \frac{\mu_t(n+s)}{\mu_t(F_s)\left[1 + \sum\limits_{\lambda=1}^{s}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda)+1}}\right] + \sum\limits_{m=1}^{s}\mu_t(m)\sum\limits_{\lambda=1}^{m}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}}} \\
&> \frac{\mu_t(n+s)}{\mu_t(F_s)\left\{1 + \frac{\varepsilon}{(1-\varepsilon)}\sum\limits_{\lambda=0}^{\infty}\left[\frac{\varepsilon}{(1-\varepsilon)^2}\right]^{\lambda}\right\} + \sum\limits_{m=1}^{s}\mu_t(m)\sum\limits_{\lambda=1}^{m}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}}} \\
&= \frac{\mu_t(n+s)}{\mu_t(F_s)\left\{1 + \frac{\varepsilon}{(1-\varepsilon)}\left[\frac{1}{1-\frac{\varepsilon}{(1-\varepsilon)^2}}\right]\right\} + \sum\limits_{m=1}^{s}\mu_t(m)\sum\limits_{\lambda=1}^{m}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}}} \\
&\equiv \frac{\mu_t(n+s)}{\mu_t(F_s)\left[1+e_1(\varepsilon)\right] + \sum\limits_{m=1}^{s}\mu_t(m)\sum\limits_{\lambda=1}^{m}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}}}
\end{aligned}$$

noting that $\lim_{\varepsilon \to 0} e_1(\varepsilon) = 0$.

Consider the last term in the denominator. For every $m = 1, \ldots, s-1$, we have the following inequality

$$\begin{aligned}
\sum_{\lambda=1}^{m}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}} &= \frac{\varepsilon}{(1-\varepsilon)^2}\sum_{\lambda=2}^{m+1}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}} \\
&< \frac{\varepsilon}{(1-\varepsilon)^2}\sum_{\lambda=1}^{m+1}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}}
\end{aligned}$$

Thus

$$\sum_{m=1}^{s}\mu_t(m)\sum_{\lambda=1}^{m}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}} < \sum_{m=1}^{s}\mu_t(m)\left[\frac{\varepsilon}{(1-\varepsilon)^2}\right]^{s-m}\sum_{\lambda=1}^{s}\frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}}$$

[9]The exact probability is $\sum\limits_{m=1}^{s}\mu_t(m)\left[\sum\limits_{\lambda=1}^{m-1}(1-\varepsilon)^{2\lambda-1}\varepsilon^{s-\lambda+1} + (1-\varepsilon)^{2(m-1)}\varepsilon^{s-m+1}\right]$

Now $\mu_t(m) \leq \frac{1}{z(1-\beta)}\mu_t(m+1)$ for every $m = 1, \ldots, s-1$, hence

$$\begin{aligned}
&\leq \mu_t(s) \sum_{m=1}^{s} \left[\frac{\varepsilon}{z(1-\beta)(1-\varepsilon)^2}\right]^{s-m} \sum_{\lambda=1}^{s} \frac{\varepsilon^{s-\lambda+1}}{(1-\varepsilon)^{2(s-\lambda+1)}} \\
&< \mu_t(s) \sum_{m=0}^{\infty} \left[\frac{\varepsilon}{z(1-\beta)(1-\varepsilon)^2}\right]^{m} \sum_{\lambda=0}^{\infty} \frac{\varepsilon^{\lambda+1}}{(1-\varepsilon)^{2(\lambda+1)}} \\
&= \mu_t(s) \sum_{m=0}^{\infty} \left[\frac{\varepsilon}{z(1-\beta)(1-\varepsilon)^2}\right]^{m} \sum_{\lambda=1}^{\infty} \left[\frac{\varepsilon}{(1-\varepsilon)^2}\right]^{\lambda} \\
&= \mu_t(s) \left[\frac{1}{1-\frac{\varepsilon}{z(1-\beta)(1-\varepsilon)^2}}\right] \left[\frac{\frac{\varepsilon}{(1-\varepsilon)^2}}{1-\frac{\varepsilon}{(1-\varepsilon)^2}}\right] \\
&< \mu_t(F_s)\frac{\beta}{z(1-\beta)} \left[\frac{1}{1-\frac{\varepsilon}{z(1-\beta)(1-\varepsilon)^2}}\right] \left[\frac{\frac{\varepsilon}{(1-\varepsilon)^2}}{1-\frac{\varepsilon}{(1-\varepsilon)^2}}\right] \\
&\equiv \mu_t(F_s)e_2(\varepsilon)
\end{aligned}$$

noting that $\lim_{\varepsilon \to 0} e_2(\varepsilon) = 0$.

Combining these results, we have for every $n \geq 1$, for all $h = (t, (C, c)^s)$ for every $s < t$

$$\tilde{\nu}_i^0(g_n|h) = \frac{\mu_t(n+s)}{\mu_t(F_s)} > \tilde{\nu}_i^\varepsilon(g_n|h) > \frac{\mu_t(n+s)}{\mu_t(F_s)\left[1+e_1(\varepsilon)+e_2(\varepsilon)\right]}$$

so that

$$\begin{aligned}
\tilde{\nu}_i^0(g_n|h) - \tilde{\nu}_i^\varepsilon(g_n|h) &< \frac{\mu_t(n+s)}{\mu_t(F_s)} - \frac{\mu_t(n+s)}{\mu_t(F_s)\left[1+e_1(\varepsilon)+e_2(\varepsilon)\right]} \\
&= \frac{\mu_t(n+s)}{\mu_t(F_s)} \left[\frac{e_1(\varepsilon)+e_2(\varepsilon)}{1+e_1(\varepsilon)+e_2(\varepsilon)}\right] \\
&< \frac{e_1(\varepsilon)+e_2(\varepsilon)}{1+e_1(\varepsilon)+e_2(\varepsilon)}
\end{aligned}$$

Since $\{g_1, g_2, \ldots\}$ is the support of both $\tilde{\nu}_i^0(\cdot|h)$ and $\tilde{\nu}_i^\varepsilon(\cdot|h)$, it follows that $\tilde{\nu}_i^\varepsilon(\cdot|h)$ converges to $\tilde{\nu}_i^0(\cdot|h)$ uniformly for every $h = (t, (C, c)^s)$ for every $s < t$.

Next consider a history for player i of the form $h_i = (t, (C, c)^{s-1}, (C, d))$ for $s < t$. Some additional notation is necessary. Let $P_s(\cdot)$ represent the probability distribution over s-length histories (i.e. over $\mathbb{N} \times \mathbb{N} \times (\{C, D\} \times \{c, d\})^{2s}$) induced by the strategy profile $\tau(\beta, \mathrm{z})$. The event that the history of signals seen by i's opponent is c^s will be denoted E. We will identify the history h_i with the event that player i's private history is h_i.

Note that $\tilde{\nu}_i^0(g_1|h_i) = 1$. We wish to show that $\tilde{\nu}_i^\varepsilon(g_1|h_i) \to 1$ uniformly for all h_i of the form in question. There are two cases in which the opponent will be playing g_1. First, the opponent's type could be $\omega_j \leq s+1$ in which case the opponent will play g_1 in accordance with $\tau(\beta, z)$. Second, the opponent's type could be $\omega_j > s+1$ but the opponent has seen an error and switched to g_1 in response. Thus, by Bayes' rule, the probability of g_1 conditional on h_i is given by

$$\tilde{\nu}_i^\varepsilon(g_1|h_i) = \frac{\sum_{m=0}^{s+1} \mu_t(m) P_s(h_i|t,m) + \mu_t(F_{s+1}) P_s(h_i \setminus E|t, F_{s+1})}{\sum_{m=0}^{\infty} \mu_t(m) P_s(h_i|t,m)}$$

Since this fraction is less than 1, we can subtract all terms $\mu_t(m)P_s(h_i|t,m)$ with $m \neq s$ from both the numerator and denominator to obtain the following inequality

$$\geq \frac{\mu_t(s)P_s(h_i|t,s) + \mu_t(F_{s+1})P_s(h_i \setminus E|t, F_{s+1})}{\mu_t(s)P_s(h_i|t,s) + \mu_t(F_{s+1})P_s(h_i|t, F_{s+1})}$$

Now noting that $P_s(h_i|\cdot) = P_s(h_i \setminus E|\cdot) + P_s(h_i \cap E|\cdot)$ we can further simplify by subtracting the terms involving $P_s(h_i \setminus E|\cdot)$ from numerator and denominator

$$\begin{aligned} &\geq \frac{\mu_t(s)P_s(h_i \cap E|t,s)}{\mu_t(s)P_s(h_i \cap E|t,s) + \mu_t(F_{s+1})P_s(h_i \cap E|t, F_{s+1})} \\ &= \frac{\mu_t(s)(1-\varepsilon)^{2s}}{\mu_t(s)(1-\varepsilon)^{2s} + \mu_t(F_{s+1})\varepsilon(1-\varepsilon)^{2s-1}} \\ &= \frac{1}{1 + \frac{\mu_t(F_{s+1})}{\mu_t(s)} \frac{\varepsilon}{1-\varepsilon}} \end{aligned}$$

Finally, noting that $\frac{\mu_t(F_{s+1})}{\mu_t(s)} < \frac{\mu_t(F_s)}{\mu_t(s)} < \frac{1-\beta}{z\beta}$, we have

$$> \frac{1}{1 + \frac{1-\beta}{z\beta} \frac{\varepsilon}{1-\varepsilon}}$$

Since this bound holds for all $s < t$, we conclude $\tilde{\nu}_i^\varepsilon(g_1|h_i) \to 1 = \tilde{\nu}_i^0(g_1|h_i)$ so that

$$\tilde{\nu}_i^\varepsilon(\cdot|h_i) \to \tilde{\nu}_i^0(\cdot|h_i)$$

uniformly for all $h_i = (t, (C,c)^{s-1}, (C,d))$ with $s < t$.

Finally consider a history h_i for player i in which i played D in the last period. Conditional on any such history, the probability that the opponent's continuation strategy is g_1 is at least the probability that the opponent correctly saw d in the last period. This probability is $(1-\varepsilon)$. Hence $\tilde{\nu}_i^\varepsilon(g_1|h_i) \to 1 = \tilde{\nu}_i^0(g_1|h_i)$ implying

$$\tilde{\nu}_i^\varepsilon(\cdot|h_i) \to \tilde{\nu}_i^0(\cdot|h_i)$$

uniformly for all such h_i.

The histories considered are the only histories for i consistent with the strategy. We have thus established the result for player i. The symmetric argument applies to player j.

QED

To this point, it has been assumed that $l \geq 1$. This implies that when a type ω_i reaches stage ω_i, he will strictly prefer to defect as the loss l from having cooperating and seeing the opponent of type $\omega_j = \omega_i$ defecting is larger than the possible gains from being able to cooperate with an opponent of $\omega_j = \omega_i + k$ for $k-1$ periods. If instead $l < 1$, then in the correlated strategy used above, type ω_i will be willing to take the chance of being defected against if his opponent's type is $\omega_j = \omega_i$, and will prefer to cooperate in hopes that the opponent's type is $\omega_j = \omega_i + k$. These strategies then fail to be an equilibrium. In this case, the correlation device μ derived from the following function f forms a correlated equilibrium:

$$f(\omega_1, \omega_2) = \begin{cases} (1-\beta)^{\frac{1}{2}(\omega_1+\omega_2-1)} & \text{if } \omega_2 \in \{\omega_1 - 1, \omega_1 + 1\} \\ z(1-\beta)^{\frac{1}{2}(\omega_1+\omega_2-1)} & \text{otherwise} \end{cases}$$

For $z = 0$, this correlation device generates two disjoint common knowledge components, one of which is illustrated in figure 5 below. For z and β close to zero and δ close to 1, it is a uniformly strict correlated equilibrium for type ω_i to play g_{ω_i}. Verifying this involves checking incentives as in Lemma 1, and is omitted. Then, just as in the case of $l \geq 1$, because all histories consistent with the strategy occur with positive probability when $\varepsilon = 0$, the beliefs after these histories change continuously as ε increases. Finally, the stationarity implies that this continuity is uniform and therefore that the equilibrium is robust.

5.3. Convergence to Grim Trigger

The correlated strategy profiles analyzed in the previous section converge to the pure grim-trigger strategies in the weak sense that behavior after every history converges to that prescribed by grim trigger. Nevertheless, there are two senses in which these strategies stay far away from the grim trigger profile, even in the limit. First, the true grim trigger strategy is never played with positive probability, and second, each player always believes that either he or his opponent will unilaterally defect after some history. In this section, I sketch how the correlated profile from the previous section can be embedded in a larger correlation device to obtain convergence in these stronger senses.

Let g_∞ denote the true grim trigger strategy. Fix values β and z for which $\tau(\beta, z)$ is a strict correlated equilibrium, and consider the following two-stage correlation device. First, a type profile $(\hat{\omega_1}, \hat{\omega_2})$ is drawn according to the correlation device μ with parameters β and z. The players are

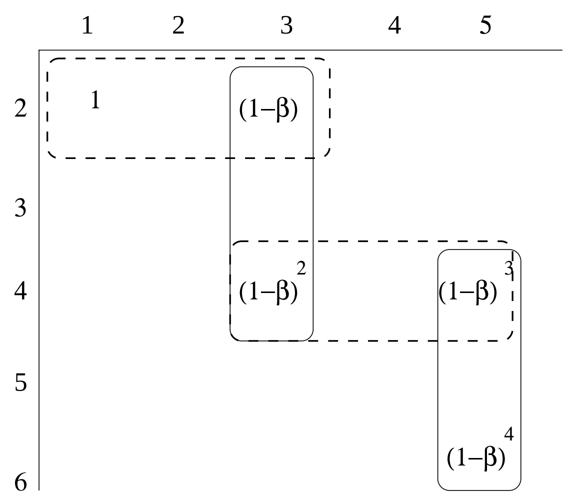

FIG. 5 Correlation device

not informed of the profile drawn. Next, a type profile is selected from the distribution represented by the following matrix, and each player is informed of his realized type.

	$\hat{\omega}_2$	g_∞
$\hat{\omega}_1$	γ	y
g_∞	y	1

Let $\tau(\gamma, y)$ be the correlated profile in which types are drawn as above and each player plays the strategy corresponding to his realized type. Suppose $\varepsilon = 0$, so that monitoring is perfect. Then for every $\gamma > 0$ there is a $y(\gamma) > 0$ such that $\tau(\gamma, y(\gamma))$ is a strict correlated equilibrium. To see this, note that for sufficiently small y, the strategies prescribed for a type $\omega_i < \infty$ remain strictly optimal. A player of type g_∞ assigns conditional probability close to 1 that his opponent's type is also g_∞. Therefore, such a player should begin by cooperating. However, there is a chance that his opponent is of type $\omega_j < \infty$ and therefore any signal of d will be interpreted as evidence that the opponent has begun defecting. Thus, type g_∞ will respond to a signal d by defecting, hence his optimal strategy is g_∞.

Just as in the previous section, since every history consistent with a player's own strategy has positive probability, and because of the structure of the conditional beliefs implied by μ, this equilibrium is robust. In particular, it remains an equilibrium for ε positive but sufficiently small. With these observations in hand, we can prove the following:

THEOREM 2. *For all sequences $\varepsilon_k \downarrow 0$, there is a sequence of correlated equilibria of $G^\infty(\delta, \varepsilon_k)$ along which the probability that the profile (g_∞, g_∞) is played converges to 1.*

Proof. Let $\varepsilon_k \downarrow 0$. Let $\gamma_n \downarrow 0$ be any sequence of positive numbers decreasing to zero. Note that (g_∞, g_∞) occurs with probability approaching 1 in any subsequence of $\tau(\gamma_n, y(\gamma_n))$.

For each n there is a $\bar{\varepsilon}_n > 0$ such that $\tau(\gamma_n, y(\gamma_n))$ is a correlated equilibrium for all $\varepsilon < \bar{\varepsilon}_n$. Define $n(k)$ to be the largest n such that $\varepsilon_k < \bar{\varepsilon}_n$. (If there is no largest n then there is a subsequence of $\bar{\varepsilon}_n$ such that $\varepsilon_k < \bar{\varepsilon}_n$ for all n. This implies that $\tau(\gamma_n, y(\gamma_n))$ is a correlated equilibrium for all $\varepsilon < \varepsilon_k$, in particular for ε_{k+n}. and we are done.) The sequence $n(k)$ is weakly increasing ($\varepsilon_{k+1} \leq \varepsilon_k < \bar{\varepsilon}_{n(k)}$ so $n(k+1) \geq n(k)$.) and unbounded (if $n(k) < N$ for all k then $\bar{\varepsilon}_{N+1} < \varepsilon_k$ for all k implying $\bar{\varepsilon}_{N+1} \leq 0$, a contradiction).

Since $\varepsilon_k < \bar{\varepsilon}_{n(k)}$, the strategy $\tau(\gamma_{n(k)}, y(\gamma_{n(k)}))$ is a correlated equilibrium for ε_k for every k, concluding the proof. QED

6. CONCLUSION

Finally, let's consider the payoff restriction $g < 1$. In either of the two correlation structures considered above (with $z = 0$), each type ω_s eventually comes to a date say $t < s$, when there are only two possible types of the opponent ω_{t+1} and ω_{s+k}. The equilibrium demands that he cooperate at such a t. Let p denote the conditional probability, having cooperated up to period t of the opponent's type being ω_{t+1} and $1 - p$ the conditional probability of ω_{t+k+1}. Defecting in period t gains $1 + g$ for sure in stage t and zero forever after. Whereas cooperating and instead defecting in stage $t + 1$ gains 1 in stage t and $(1 - p)(1 + g)$ in the next stage. Given discounting, if $g \geq 1$, then the latter can exceed the former only if $p < 1/2$. But it is impossible for infinitely many types ω_s to assign probability exceeding $1/2$ to higher types for the opponent. Thus, these equilibria require $g < 1$.

The crucial feature of these equilibria is the unilateral deviation that occurs on the equilibrium path in every period. Whether there exist other approaches to the construction of robust correlated trigger-strategy equiilbria when $g < 1$ remains an open question.

REFERENCES

BAGWELL, K. (1995): "Commitment and Observability in Games," *Games and Economic Behavior*, 8, 271–280.

BHASKAR, V. (1998): "Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems," *Review of Economics Studies*, 65, 135–149.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 102(1), 40–69.

Bhaskar, V., and E. van Damme (2002): "Moral Hazard and Private Monitoring," *Journal of Economic Theory*, 102(1), 16–39.

Ely, J. C., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102(1), 84–105.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," *Econometrica*, 54, 533–554.

Mailath, G. J., and S. Morris (2002): "Repeated Games with Almost Public Monitoring," *Journal of Economic Theory*, 102(1), 189–228.

Matsushima, H. (1991): "On the Theory of Repeated Games with Private Information, part I: Anti-Folk Theorem without Communication," *Economics Letters*, 35, 253–256.

Piccione, M. (2002): "The Repeated Prisoners' Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1), 70–83.

Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review*, 79, 385–391.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 76(2), 345–361.