The Value of Linking: Efficiency and Public Punishment*

Jimmy Chan[†] Wenzhang Zhang[‡]

September 4, 2015

Abstract

We consider repeated games with mediation and side-payments. We show that any punishment scheme can be divided into a public part and a private part. While private punishment can be reduced by linking, public punishment must be carried out in equilibrium. The total average equilibrium payoff is bounded from above by the payoff from enforcing the correlated action profile with the highest total stage-game payoff net of public punishment. The bound is tight. For any ϵ there exists an equilibrium with total average payoff ϵ close to the upper bound when the players are sufficiently patient. Our results extend the insights of Abreu, Milgrom, and Pearce (1991) and incorporate a number of existing results in the literature.

1 Introduction

In recent years the theory of repeated games has made substantial progress in understanding how disperse, noisy information can be utilized to enforce cooperative behavior in long-run relationships. An important insight that has emerged from

*Preliminary and Incomplete.

[†]Fudan University. Email: jimmy.hing.chan@gmail.com

[‡]Shanghai University of Finance and Economics. Email:wenzhang.zhang@gmail.com

this literature is that optimal incentives are often non-linear. In particular, when monitoring is non-public, linking payoff decisions across periods may reduce the cost of imperfect monitoring.

The idea is first introduced by Abreu, Milgrom, and Pearce (1991). They consider a two-person partnership game. To induce cooperation both players must be punished at a cost when a bad signal occurs. Abreu, Milgrom, and Pearce (1991) show that if the signals are publicly observed immediately at the end of each period, the costly punishment must be carried out immediately when a bad signal occurs. However, if the signals are observed with a lag, the players can delay the punishment and use the same punishment to induce cooperation in multiple periods. Intuitively, not observing the signals immediately reduce the number of incentive compatibility constraints that an equilibrium must satisfy.

In Abreu, Milgrom, and Pearce (1991) the timing of information is exogenous. Subsequent research has shown that in games of private monitoring a similar result may be obtained when players delay revealing their private signals endogenously. The key issue is to ensure that players have the right incentives to report and do not learn “too much” from their own private signals. In this paper we introduce a new approach that unify many of the existing results. Our approach is closely related to the original insights of Abreu, Milgrom, and Pearce (1991). We show that any punishment scheme can be divided into a public part and a private part. Only the private part can reduced through a linking mechanism. The public part must be carried out in equilibrium.

In our model the players can employ a mediator to implement a correlated stage-game outcome, and they can exchange side-payments at the end of each period. The value of mediation has been demonstrated by Rahman (2014). He shows that in a repeated game with public monitoring the players may use a correlated strategy to achieve approximate efficiency as the discount factor goes to one even when the public information arrives continuously. Without a mediator the players may still correlate their actions through private signals in the past (Kandori and Obara, 2006), but this sort of private-strategy equilibrium is hard to analyze. While uncommon in repeated games, side-payments are natural in many economic problems. Allowing side-payments allow us to avoid the difficulty of enforcing an outcome with rewards (rather than punishments). This enables us establish a tight bound on the equilibrium efficiency loss. Existing results, by contrast, only identify conditions under which the efficiency loss can be made arbitrarily small.

A key question of our analysis is what is public when information are private? In our model in each period each player receives a private recommendation from a mediator before choosing an action, and observes a private signal after the actions are chosen. The mediator makes recommendations according to a correlated action profile. The correlated action profile and the signal distribution jointly induce a distribution over actions and signals. At the end of each period, the players form beliefs over actions and signals on the basis of the recommendation and signal that he receives.

A set of action-signal profiles is self-evident if given the realization of any member of the set it is common belief among players that the realized action-signal profile is contained in the set.[1] The set of action-signals profiles can be partitioned into self-evident sets. We argue that self-evident sets are public as the players beliefs would not the change if, before observing their private signals, they were first publicly informed of the self-evident set that contains the realized action-signal profile. We show that of any punishment schemes the part that depends on these "public" signals must be carried out in equilibrium. Thus, an upper bound of the total average equilibrium payoff that can be achieved by enforcing the correlated action profile that has the highest total stage-game payoff net of public punishment. Since non-public punishment can be eliminated through linking, for any ϵ there exists an equilibrium with total average payoff arbitrarily ϵ close to the upper bound when the players are sufficiently patient.

An immediate corollary of the last result is that approximate efficiency can be attained when an efficient pure-action profile can be enforced without public pun-

[1]Given the realization of a specific signal profile, a set of signal profiles is common belief among the players if every player believes that the realized profile belongs to the set with probability one, and every players believes with probability one every player believes the the realized profile belongs to the set with probability one, and so forth.

ishment. Previous results on approximate efficiency are based on two approaches. One relies on linking. Compte (1998) assumes the players' private signals are independent. Obara (2009), Zheng (2008) and Chan and Zhang (forthcoming) consider correlated signals with full support. Another approach is to ensure that a non-deviator can always be identified and rewarded to keep the total punishment zero (Fudenberg, Levine, and Maskin, 1994, Kandori and Matsushima, 1998). Our result incorporates both approaches. When the signal distribution has full support, there is only one self-evident set that contains all signal profiles. In this case any enforceable action profile can be enforced without public punishment. Alternatively, when every signal profile is self-evident, no public punishment is possible only when a non-deviator can be identified. In generally, our result says that an efficient outcome can be enforced with no public punishment if any unilateral deviation that cannot be deterred using the signals within a self-evident set must lead to a distribution of self-evident sets that cannot be caused by another player.

2 Model

We consider a mediated repeated game with communication and side-payments that we denote by $\Gamma^{\infty}(G, B, \chi)$. Time periods are denoted by $t = 0, 1, 2, \ldots$. The players are denoted by $i = 1, 2, \ldots, n$. Let $\mathcal{N} = \{1, 2, \ldots, n\}$.

The events in period 0 unfold as follows. Nature draws $\beta = (\beta_1, \ldots, \beta_n)$ from a countable set $B = B_1 \times \cdots \times B_n$ according to a distribution χ. Each player i observes β_i and sends a public message $m_i \in M_i$ to the other players. We assume that M_i is sufficiently rich that player i can convey any private information during the the repeated game. Each player i then simultaneously makes a publicly observable side-payment τ_{ij} to each player j.[2] Finally, the players observe ζ, the outcome of a public randomization device that is uniformly distributed between 0 and 1.

[2]We include τ_{ii}, player i's payment to himself, to simplify notation. Throughout, we set τ_{ii} to zero.

In each period $t = 1, 2, \ldots,$ the players play the following stage game G. Let $A = A_1 \times ... \times A_n$ denote a finite set of action profile. First, a mediator first picks a distribution $\eta \in \Delta(A)$, draws $\widetilde{a} = (\widetilde{a}_1, ..., \widetilde{a}_n) \in A$ according to η, and then privately informs each player $i \in \mathcal{N}$ of $\widetilde{a}_i$. After observing $\widetilde{a}_1$, each player $i \in \mathcal{N}$ simultaneously chooses a private action a_i from A_i. Nature then draws $y = (y_1, \ldots, y_n)$ from a finite set $Y = Y_1 \times \cdots \times Y_n$ according to a distribution $p(y|a)$. Each player i privately observes y_i and sends a public message $m_i \in M_i$ to the other players. The mediator publicly reports $\kappa \in (A \cup \{\emptyset\})$. Each player i then simultaneously makes a publicly observable side-payment τ_{ij} to each player j. Finally, the players observe the outcome of a public randomization device ϕ.

Let a_{-i} and y_{-i} denote a minus a_i and y minus y_i, respectively.[3] The marginal probabilities of y_{-i}, y_i and (y_i, y_j) are denoted respectively by $p_{-i}(y_{-i}|a)$, $p_i\,(y_i|a)$ and $p_{ij}(y_i, y_j|a)$, and the marginal probabilities of y_{-i} and y_j, conditional on a and y_i, are denoted respectively by $p_{-i}(y_{-i}|a, y_i)$ and $p_j\,(y_j|a, y_i)$. To simplify exposition, we assume that the support of p is independent of a and use Y^* to denote the common support. Beyond common support we do not impose any other restrictions on p.

Player i's stage-game payoff, denoted by $r_i\,(a_i, y_i)$, depends only on his own action and signal. The actions of the other players affect player i's payoff through their effects on p. Player i's expected stage-game payoff conditional on a is

$$g_i(a) \equiv \sum_{y_i \in Y_i} r_i(a_i, y_i) p_i(y_i|a).$$

To save notation for any $\eta \in \Delta(A)$ we let

$$g_i(\eta) \equiv \sum_{a \in A} g(a)\eta(a).$$

The mediator is indifferent over all outcomes.

For each variable x, we use x_t to denote the period-t value of x, and x^t to denote the history of x up to period t. For example, $m^t = (m_0, m_1, \ldots, m_t)$ is the history

[3]For any variable x_i we use x to denote $(x_1, x_2, \ldots, x_n)$ and x_{-i} to denote x with the i-th element x_i deleted.

of players' message-report profiles up to period t. We use $x_{i,t}$ to denote period-t value of x_i. Because it will usually be clear from the context that whether, say, a_1, refers to the action of player 1 or the action profile chosen in period 1, we do not introduce extra notations to distinguish the two. We will state the meaning explicitly when confusion may arise.

A history of the game is an infinite sequence

$$h^\infty \equiv (\beta, m_0, \tau_0, \zeta_0, \widetilde{a}_1, a_1, y_1, m_1, \kappa_1, \tau_1, \zeta_1, \ldots, \widetilde{a}_t, a_t, y_t, m_t, \kappa_t, \tau_t, \zeta_t, \ldots).$$

The players discount future payoffs by a common discount factor $\delta < 1$. Given a history h^∞, player i's average repeated-game payoff is

$$u_{i,0}(h^\infty) \equiv (1-\delta)\left(\sum_{j=1}^{n}(\tau_{ji,0} - \tau_{ij,0}) + \sum_{t=1}^{\infty}\delta^t\left(r_i(a_{i,t}, y_{i,t}) + \sum_{j=1}^{n}(\tau_{ji,t} - \tau_{ij,t})\right)\right).$$

Similarly, for each $s \geq 1$, player i's average payoff from the beginning of period s onward, given h^∞, is

$$u_{i,s}(h^\infty) \equiv (1-\delta)\sum_{t=s}^{\infty}\delta^t\left(r_i(a_{i,t}, y_{i,t}) + \sum_{j=1}^{n}(\tau_{ji,t} - \tau_{ij,t})\right).$$

At the beginning of each period t, both the mediator and all players have observed a public history h_t^{pub} that consists of the signal reports of the players, public message of the mediator, side-payments, and outcomes of the public randomization device in the previous t periods. Let H_t^{pub} denote the set of period-t public histories, and let

$$H_t^{pub+} \equiv H_t^{pub} \times M_1 \times \cdots \times M_n$$

denote the set of period-t public histories that includes the players' reports in period t. A strategy of the mediator is a function $\sigma_m = (\sigma_{m1}, \sigma_{m2})$ where σ_{m1} is a selection strategy that maps each history in $\cup_{t\geq 1} H_t^{pub}$ into a distribution η in $\Delta(A)$, and σ_{m2} is a reporting strategy that maps each history in each history in $\cup_{t\geq 1} H_t^{pub+}$ into a report $\kappa \in (A \cup \{\emptyset\})$.

In addition to the public history, player i has observed a private history $h_{i,t}^{pri}$ that consists of his observation of the correlating device β_i in period 0, the private recommendation from the mediator, and his actions and signals in the previous $(t-1)$ periods. We use $h_{i,t} = \left(h_t^{pub}, h_{i,t}^{pri}\right)$ to denote the history, both public and private, that player i observes at the beginning of period t. Let $H_{i,t}^0$ denote the set of histories $h_{i,t}$ for player i at the beginning of period $t \geq 1$, and

$$H_{i,t}^1 \equiv H_{i,t}^0 \times A_i$$

the set of histories $h_{i,t}$ for player i after he receives the private recommendation from the mediator in period $t \geq 1$. Let

$$H_{i,t}^2 \equiv \begin{cases} B_i, & \text{if } t = 0; \\ H_{i,t}^1 \times A_i \times Y_i, & \text{if } t \geq 1. \end{cases}$$

denote the set of possible histories for player i in the middle of period t after observing β_i (if $t = 0$) or after choosing $a_{i,t}$ and observing $y_{i,t}$ (if $t \geq 1$), and let

$$H_{i,t}^3 \equiv H_{i,t}^2 \times M_1 \times \cdots \times M_n \times (A \cup \{\emptyset\})$$

denote the set of possible histories for player i in the middle of period t after observing the message profile m_t and mediator's report κ_t. Let

$$H_{i,t} = H_{i,t}^1 \cup H_{i,t}^2 \cup H_{i,t}^3$$

denote the set of possible histories for player i during period t after which he needs to make a decision. A generic element of $H_{i,t}$ will be denoted by $\varphi_{i,t}$.

A pure strategy $\sigma_i = (\alpha_i, \rho_i, b_i)$ for player i consists of three components: an action strategy α_i that maps each history in $\cup_{t\geq 1} H_{i,t}^1$ into an action in A_i, a reporting strategy ρ_i that maps each history in $\cup_{t\geq 0} H_{i,t}^2$ into a message in M_i, and a transfer strategy $b_i = (b_{i1}, b_{i2}, \ldots, b_{in})$ that maps each history in $\cup_{t\geq 0} H_{i,t}^3$ into an n-vector of nonnegative real numbers. Let Σ_i denote the set of all pure strategies σ_i for player i and let $\sigma = (\sigma_1, \ldots, \sigma_n)$ denote a strategy profile.

A system of beliefs μ specifies, for each $i \in \mathcal{N}$ and each $t \geq 0$, a probability distribution $\mu_{\varphi_{i,t}}(\cdot)$ of $\left(\beta_{-i}, \widetilde{a}_{-i}^t, a_{-i}^{t-1}, y_{-i}^{t-1}\right)$ for each $\varphi_{i,t} \in H_{i,t}^1$, and a probability distribution $\mu_{\varphi_{i,t}}(\cdot)$ of $\left(\beta_{-i}, \widetilde{a}_{-i}^t, a_{-i}^t, y_{-i}^t\right)$ for each $\varphi_{i,t} \in H_{i,t}^2 \cup H_{i,t}^3$. An assessment

(σ, μ) consists of a strategy profile and a system of beliefs. Given any assessment (σ, μ) and any history $\varphi_{i,t} \in H_{i,t}$, we use

$$v_{i,t}(\sigma, \mu, \varphi_{i,t}) \equiv E\left[u_{i,t}(h^\infty) \middle| \sigma, \mu, \varphi_{i,t}\right]$$

to denote the expected value of player i's payoff from period t onward, where the expectation is taken over the distribution of histories h^∞ induced by σ, conditional on $\varphi_{i,t}$ and the belief $\mu_{\varphi_{i,t}}(\cdot)$. Write $v_i(\sigma, \mu)$ for $v_{i,0}(\sigma, \mu, \varphi_{i,0})$.

An assessment (σ_m, σ, μ) is a perfect Bayesian equilibrium if the following two conditions hold.

- For all $i \in \mathcal{N}$, $t \geq 0$, and $\varphi_{i,t} \in H_{i,t}$ such that $\Pr(\varphi_{i,t}|\sigma) > 0$, the belief $\mu_{\varphi_{i,t}}(\cdot)$ is derived from (σ_m, σ) using Bayes' rule.
- For all $i \in \mathcal{N}$, $t \geq 0$, $\varphi_{i,t} \in H_{i,t}$, and $\sigma_i' \in \Sigma_i$,

$$v_{i,t}(\sigma, \mu, \varphi_{i,t}) \geq v_{i,t}(\sigma_i', \sigma_{-i}, \mu, \varphi_{i,t}). \tag{1}$$

Note that for $\varphi_{i,t} \in H_{i,t}^2 \cup H_{i,t}^3$, (1) is equivalent to

$$E\left[(1-\delta)\delta^t \sum_{j=1}^n (\tau_{ji,t} - \tau_{ij,t}) + u_{i,t+1}(h^\infty) \middle| \sigma, \mu, \varphi_{i,t}\right]$$
$$\geq E\left[(1-\delta)\delta^t \sum_{j=1}^n (\tau_{ji,t} - \tau_{ij,t}) + u_{i,t+1}(h^\infty) \middle| \sigma_i', \sigma_{-i}, \mu, \varphi_{i,t}\right].$$

Hence the second condition implies sequential rationality.

3 Public Punishment

Let G_0 denote the stage game G without the last two steps (i.e., making side-payments and observing the outcome of the public randomization device). Suppose an mechanism designer tries to induce the players to choose a correlated strategy $\eta \in \Delta(A)$ through a transfer scheme. At the end of G_0, each player i is paid, in additional to his stage-game payoff, a transfer that depends on the players' signal report and the mediator's recommendation. Formally, let $w_i : A \times Y \to \Re$ denote player i's transfer. We call a transfer scheme, $w = (w_1, ..., w_n)$, a punishment scheme if $\sum_i w_i (\widetilde{a}, y) \leq 0$ for all $(\widetilde{a}, y) \in A \times Y$.

A pure strategy for player i in this extended game consists of an action strategy α_i that maps each private recommendation into an action, and a reporting strategy ρ_i that maps each y_i into a message in Y_i. Let Ψ_i the set of reporting strategies. We say that a punishment scheme $w = (w_1, ..., w_n)$ enforces η if it is a Nash equilibrium for the players to follow the recommendations and report their signal truthfully when the mediator selects $\widetilde{a}$ according to η.

Definition 1. A transfer function profile $w = (w_1, \ldots, w_n)$ enforces η if, for each player i and each $\widetilde{a}_i \in supp\,(\eta_i)$, $a_i' \in A_i$ and $\rho_i' \in \Psi_i$,

$$\sum_{\widetilde{a}_{-i}} \left(g_i\left(\widetilde{a}\right) + E_y\left[w_i\left(\widetilde{a}, y\right)|\widetilde{a}\right]\right) \eta_{-i}\left(\widetilde{a}_{-i}|\widetilde{a}_i\right)$$
$$\geq \sum_{\widetilde{a}_{-i}} \left(g_i\left(a_i', \widetilde{a}_{-i}\right) + E_y\left[w_i\left(\widetilde{a}, \rho_i\left(y_i\right), y_{-i}\right)|a_i', \widetilde{a}_{-i}\right]\right) \eta_{-i}\left(\widetilde{a}_{-i}|\widetilde{a}_i\right). \quad (2)$$

The enforcement is strict if strict inequality holds in (**??**) for all $a_i' \neq a_i$.

Let $W\left(\eta\right)$ denote the set of punishment schemes that strictly enforces η. For any $w \in W\left(\eta\right)$ let

$$L(w, \eta) = \sum_{i=1}^{n} \sum_{(a,y) \in A \times Y} w_i(a, y)\mu(a, y)$$

denote the total expected transfer.

To introduce the concept of public punishment, we need to first define what is "public." Fix the mediator strategy η and assume the players follow the recommendations of the mediator. Let

$$\mu\left(\widetilde{a}, y\right) \equiv p\left(y|\widetilde{a}\right) \eta\left(\widetilde{a}\right)$$

be the distribution of $(\widetilde{a}, y)$ induced by η and p, and

$$\left(A \times Y\right)\left(\eta\right) \equiv \left\{\left(\widetilde{a}, y\right) \mid \mu\left(\widetilde{a}, y\right) > 0\right\}$$

the set of $(\widetilde{a}, y)$ that is possible under η. Let $(A_i \times Y_i)(\eta)$ denote the projection of $(A \times Y)(\eta)$ on $A_i \times Y_i$. For any $i \in \mathcal{N}$, let P_i denote a partitional information function of $(A \times Y)(\eta)$ such that for each $(\widetilde{a}'_i, y'_i) \in (A_i \times Y_i)(\eta)$

$$P_i(\widetilde{a}'_i, y'_i) = \{(\widetilde{a}'_i, \widetilde{a}_{-i}, y'_i, y_{-i}) \in (A \times Y)(\eta)\}$$

denote the set of recommendation and signal profiles that player i believes is possible conditional on $(\widetilde{a}'_i, y'_i)$.

Let P denote the meet (i.e., the least common coarsening) of $\{P_i\}_{i=1}^n$. P has two important properties. First, for any $\omega \in P$, $i \in \mathcal{N}$ and $(\widetilde{a}'_i, y'_i) \in (A_i \times Y_i)(\eta)$, $\mu(.|(\widetilde{a}'_i, y'_i), \omega) = \mu(.|(\widetilde{a}'_i, y'_i))$. Second, for any $\omega \in P$, and $x \subset \omega$, there exists $i \in \mathcal{N}$ and $(\widetilde{a}'_i, y'_i) \in (A_i \times Y_i)(\eta)$ such that $\mu(x|(\widetilde{a}'_i, y'_i)) > 0$ and $\mu(.|(\widetilde{a}'_i, y'_i), x) \neq \mu(.|(\widetilde{a}'_i, y'_i))$. In the terminology of interactive epistemology, any element of P is self-evident and any proper subset of any element of P is not (Osborne and Rubinstein Ch.5).[4] We call ω a "public signal" because each player i "observes" ω in the sense he can always infer from $(\widetilde{a}_i, y_i)$ which $\omega \in P$ contains $(\widetilde{a}, y)$, and each player knows that the other players also observe ω, and so on.

For any $w \in W(\eta)$ we call

$$L^{pub}(w, \eta) = L(w, \eta) - \max_{\omega \in P} \sum_i \sum_{\widetilde{a}, y} w_i(\widetilde{a}, y) \mu(\widetilde{a}, y|\omega)$$

the total public transfer. We are interested in the minimum amount of public punishment (i.e., the absolute value of the total public transfer) needed to strictly enforce η. Let

$$Q(\eta) \equiv conv\{\mu(\cdot|\eta, \omega) \mid \omega \in P\}$$

denote the set of distributions that have the same conditional distributions as μ, and let $s_i : A_i \to \Delta(A_i \times \Psi_i)$ denote a mixed strategy in G_0 where $s_i(\widetilde{a}_i, a_i, \rho_i)$ is the probability of choosing (a_i, ρ_i) after receiving the recommendation $\widetilde{a}_i$. The

[4]A subset E of of $(A \times Y)(\eta)$ is an event. Player i "knows" an event when he observes $(\widetilde{a}_i, y_i)$ if $P_i(\widetilde{a}_i, y_i) \subseteq E$. An event E is a common belief at $(\widetilde{a}, y)$ if every player i knows E, knows every player $j \neq i$ knows E, and so on when the mediator recommends $\widetilde{a}$ and the signal profile y occurs. An event E is self-evident if it is common belief at any $(\widetilde{a}, y) \in E$.

distribution of $(\widetilde{a}, y)$ induced by $\left(s_i, \alpha^*_{-i}\right)$ is π^{s_i}. For each $(\widetilde{a}, y) \in A \times Y$,

$$\pi^{s_i}(\widetilde{a}, y) = \eta(\widetilde{a}) \sum_{a_i, \rho_i} s_i(\widetilde{a}_i, a_i, \rho_i) \sum_{y'_i \in \rho_i^{-1}(y_i)} p(y_{-i}, y'_i | \widetilde{a}_{-i}, a_i).$$

Since P is a partition, for any $\pi^{s_i} \in Q(\eta)$ there is a unique $\widetilde{\nu} : P \to [0, 1]$ such that $\sum_{\omega \in P} \widetilde{\nu}(\omega) = 1$ and

$$\pi^{s_i}(\cdot) = \sum_{\omega \in P} \widetilde{\nu}(\omega) \mu(\cdot | \omega).$$

Let

$$\gamma(s_i) \equiv \min_{\omega \in P} \left(\frac{\mu(\omega)}{\widetilde{\nu}(\omega) - \mu(\omega)} \right).$$

The following proposition establishes the minimum public punishment needed to enforce η.

Proposition 1. *For any strictly enforceable η, $L^*(\eta) \equiv \sup_{w \in W^*(\eta)} L(w, \eta)$ is equal to*

$$\inf_{(s_1, \dots, s_n)} \gamma(s_1) \left[\sum_{i=1}^{n} \sum_{\widetilde{a} \in A} \eta(\widetilde{a}) \sum_{a_i, \rho_i} s_i(\widetilde{a}_i, a_i, \rho_i) \left(g_i(\widetilde{a}) - g_i(\widetilde{a}_{-i}, a_i) \right) \right] \quad (3)$$

$$s.t. \pi^{s_1} = \cdots = \pi^{s_n} \in Q(\eta). \quad (4)$$

To illustrate Proposition 1, consider a Noisy Prisoners' Dilemma game. Each player chooses C or D and then observes a public signal that is either H or L. The stage-game payoffs and signal distribution are given in Figure 1. It is assumed that $p > q$ and $h, d > 0$.

Actions

	C	D
C	$1, 1$	$-h, 1 + d$
D	$1 + d, -h$	$0, 0$

Public Signal Dist.

	H	L
CC	p	$1 - p$
CD	q	$1 - q$

When $\eta(CC) = 1$. the information structure is

$$P_1(C, H) = \{H\}, P_1(C, L) = \{L\};$$
$$P_2(C, H) = \{H\}, P_2(C, L) = \{L\};$$
$$P = \{H\}, \{L\}.$$

In this case each element of P is a singleton. $Q\left(CC\right)$ is therefore any distribution over H and L. To ensure that player i prefers C to D, w_i must satisfy the incentive-compatibility constraint that

$$\left(w_i\left(H\right) - w_i\left(L\right)\right) \geq \frac{d}{\left(p-q\right)}. \tag{5}$$

Since the game is symmetric, public punishment requires that

$$w_i\left(H\right), w_i\left(L\right) \leq 0. \tag{6}$$

Maximizing the total expected transfer

$$\sum_{i=1,2}\left(pw_i\left(H\right) + \left(1-p\right)w_i\left(L\right)\right)$$

subject to (5) and (6) yields $w_i\left(H\right) = 0$ and $w_i\left(L\right) = -d/\left(p-q\right)$ for $i = 1,2$. The maximum total expected transfer is $2d\left(1-p\right)/\left(p-q\right)$. The factor $\left(1-p\right)/\left(p-q\right)$ measures the efficacy of the transfer scheme. This factor is s-maller if the deviation leads to a larger change in probability distribution (i.e., a bigger $p-q$) or if the "bad" signal L is unlikely to occur when the players do not deviate. Since each player can unilaterally deviate to D and gain d, the total deviation gain that the scheme needs to deter is $2d$. The total expected transfer is the product of the total deviation gain and the efficacy of the transfer scheme.

In general there are multiple ways to deviate. Proposition 1 says that we only need to worry about a deviating strategy s_i if $\pi^{s_i} \in Q\left(\eta\right)$ and there exists s_j such that $\pi^{s_j} = \pi^{s_i}$ for each $j \neq i$. If a deviation $\pi^{s_i} \notin Q\left(\eta\right)$, then s_i can be deterred by using private signals within some $\omega \in P$. If there is no s_j such that $\pi^{s_j} = \pi^{s_i}$, then player j cannot cause π^{s_i}, and we can punish player i and reward player j when π^{s_i} "occurs" to keep the total punishment zero.

For any $s_1, ..., s_n$ that satisfy (4), the minimum total punishment needed to deter each player i from choosing s_i is

$$\gamma(s_1)\left[\sum_{i=1}^{n}\sum_{\widetilde{a}\in A}\eta\left(\widetilde{a}\right)\sum_{a_i,\rho_i}s_i\left(\widetilde{a}_i, a_i, \rho_i\right)\left(g_i\left(\widetilde{a}\right) - g_i\left(\widetilde{a}_{-i}, a_i\right)\right)\right].$$

The term inside the square bracket is the total deviation gains, and $\gamma(s_1)$ is the efficacy of the best statistical test (i.e., the counterpart of $(1-p)/(p-q)$). Proposition 1 says that the total public punishment needed to enforce a distribution is the public punishment needed to deter the "worst" deviation.

4 Maximum Equilibrium Payoff

Proposition 2. *For any $\delta \in (0,1)$, the total average payoff of any perfect action-public equilibrium of $\Gamma^{\infty}(B,\chi,G,\delta)$ in which the action strategies are pure is less than*

$$S^{*} \equiv \sup_{\eta} \left(\sum_{\widetilde{a}} \sum_{i=1}^{n} g_i(\widetilde{a})\, \eta(\widetilde{a}) + L^{*}(\eta) \right).$$

Furthermore, any $\epsilon > 0$, there exists $\bar{\bar{\delta}} < 1$ such that for each $\delta > \bar{\bar{\delta}}$, there is a perfect action-public equilibrium of $\Gamma^{\infty}(B,\chi,G,\delta)$ with total average equilibrium payoff greater than $S^{*} - \epsilon$.

Part 1 of Proposition 2 provides an upper bound to the total average equilibrium payoff. Part 2 of Theorem 2 shows that the bound is tight when the players are sufficiently patient.

Necessity. Part 1 generalizes the inefficiency result of Abreu, Milgrom, and Pearce (1991). We illustrate the main idea with a simple case. Fix δ. Assume that all signals are public and there is an equilibrium that involves only pure-action profiles and attains maximum total equilibrium payoff. Since both signals and actions are public. Let $v^{\max}$ be the maximum total equilibrium payoff and σ an equilibrium that attains $v^{\max}$. Suppose a is chosen in the first period in σ. Since the equilibrium action is pure and all signals are public, there is no private information on the equilibrium path, and the continuation strategies depend only on the public signal y. Let $v_i(y,\sigma)$ denote player i's equilibrium continuation payoffs after y, and write $w_i(y)$ for $(v_i(y,\sigma) - v^{\max})\,\delta/(1-\delta)$. By definition,

$$v^{\max} = \sum_i \left((1-\delta)\, g_i(a) + \delta \sum_y v_i(y,\sigma)\, p(y|a) \right), \tag{7}$$

$$= \sum_i \left((1-\delta)\, g_i(a) + \delta v^{\max} + (1-\delta) \sum_y w_i(y,\sigma)\, p(y|a) \right) \tag{8}$$

Subtract $\delta v^{\max}$ from both sides, and divide both sides by $(1-\delta)$. We obtain

$$v^{\max} = \left(\sum_i g_i(a) + \sum_y w_i(y)\, p(y|a) \right). \tag{9}$$

Since it is optimal for each player i to choose a_i in the first period, $w = (w_1, ..., w_n)$ enforces a. Moreover, as the continuation game following any y is itself a repeated game, for any y

$$\sum_i w_i(y) = \frac{\delta}{1-\delta} \left(\sum_i v_i(y,\sigma) - v^{\max} \right) \le 0.$$

It follows that w is a public-punishment scheme and, hence,

$$v^{\max} \le \sum_i (1-\delta)\, g_i(a) + L(a,G) \le S^*. \tag{10}$$

When signals are private or equilibrium action profile is correlated, the continuation game after the players observe their first period signals is not itself a repeated game as each player i has private information about $(\widetilde{a}_i, y_i)$. The total continuation payoff conditional on $(\widetilde{a}, y)$ is, therefore, not bounded by the maximum equilibrium payoff. To apply the recursive argument, we consider a modified game where in each period the players observe publicly $\omega \in P$ before learning their private signals. Since ω is self-evident, the modified game will have the same set of equilibria as the original game. Furthermore, since the continuation game following each ω is itself a mediated repeated game, the total continuation payoff conditional on ω is bounded by the upper bound of the original game.

We prove part 2 of Theorem 2 by constructing a T-period mechanism that strictly enforce η in a T-period repetition of G_0 with expected total transfer ϵ

close to $L(\eta)$. Intuitively, almost all non-public punishment can be eliminated by linking when the players are sufficiently patient. With the mechanism, it is straightforward to construct the required perfect action-public equilibrium along the lines of Chan and Zhang (forthcoming).

References

Abreu, Dilip, Paul Milgrom, and David Pearce, 1991, Information and timing in repeated partnerships, *Econometrica* 59, 1713–1733.

Chan, Jimmy, and Wenzhang Zhang, forthcoming, Approximate efficiency in repeated games with side-payments and correlated signals, *Theoretical Economics.*

Compte, Olivier, 1998, Communication in repeated games with imperfect private monitoring, *Econometrica* 66, 597–626.

Fudenberg, Drew, David Levine, and Eric Maskin, 1994, The folk theorem with imperfect public information, *Econometrica* 62, 997–1039.

Kandori, Michihiro, and Hitoshi Matsushima, 1998, Private observation, communication and collusion, *Econometrica* 66, 627–652.

Kandori, Michihiro, and Ichiro Obara, 2006, Efficiency in repeated games revisited: The role of private strategies, *Econometrica* 74, 499–519.

Obara, Ichiro, 2009, Folk theorem with communication, *Journal of Economic Theory* 144, 120–134.

Rahman, David, 2014, The dilemma of the cypress and the oak tree, Discussion paper.

Zheng, Bingyong, 2008, Approximate efficiency in repeated games with correlated private signals, *Games and Economic Behavior* 63, 406 – 416.