Communication in a Dynamic Prisoner's Dilemma with Incomplete Information*

Yu Awaya†

April 11, 2013

Abstract

This paper provides a model in which communication is *necessary* in order to achieve cooperative outcomes in a long-term relationship. The model is a dynamic prisoner's dilemma with incomplete information about payoffs. The payoffs are private information and stochastically evolve over time. I study two situations. In one, players play simultaneously in every stage, knowing their own types. In the other, players exchange cheap talk messages after knowing their own types but prior to play. I show that there exists a nearly efficient payoff vector that is achieved as an equilibrium outcome when communication is possible and players are patient, but cannot be achieved without communication no matter how patient the players are.

1 Introduction

According to the Bible, a united humanity started to build the Tower of Babel "whose top may reach unto heaven ..." This attempt to reach heaven displeased God greatly He said: "They are one people and have one language, and nothing will be withheld from them which they purpose to do." So God thwarted humanity's efforts by "confound[ing] the language of all the Earth". By separating the languages, God disrupted communication among the mortals preventing them from succeeding in their cooperative endeavor.

At a more earthly level, antitrust law deems certain kinds of communication between competing firms to be illegal, again based on the premise

*I am most grateful to Vijay Krishna for his guidance and encouragement. I also thank Kalyan Chatterjee, Ed Green, Jim Jordan, Sho Miyamoto, John Stachurski, Tadashi Sekiguchi and Neil Wallace for their helpful comments.

†Department of Economics, Penn State University, yxa120@psu.edu.

that communication facilitates collusion. God and policymakers share the common and intuitive idea that communication is *necessary* in order to cooperate/collude. The literature on repeated games, however, does not provide theoretical support for this idea. The folk theorem for repeated games (Fudenberg and Maskin (1986)) shows that cooperation can be achieved without any communication among the players. The same is true for repeated games with imperfect public monitoring (Fudenberg, Levine and Maskin (1994)). With private monitoring, Compte (1998) and Kandori and Matsushima (1998) establish folk theorems in situations in which players can communicate with each other. But recently, Sugaya (2011) has shown that even in games with private monitoring, communication can be dispensed with altogether. One lesson of these results is that communication is not necessary for cooperation.

The results mentioned above concern games of complete information. But even with incomplete information, there are folk theorems which do not make use of any communication. Specifically, in games in which types are drawn only once (and so are "perfectly persistent"), Fudenberg and Yamamoto (2010, 2011) and Yamamoto (2012) prove such a result under various assumptions.[1]

In a Bertrand oligopoly game in which types are independently drawn in every period (and so are not persistent), Athey and Bagwell (2001) informally claim that communication improves equilibrium outcome. But they consider only public perfect equilibrium (PPE) and Hörner and Jamison (2007) have shown that once the restriction to PPE is removed, collusive outcomes can be achieved without any communication. The overall message is again that communication has little to do with cooperation.

This paper studies a model with evolving incomplete information in which *communication strictly improves equilibrium payoffs*. Specifically, there are (symmetric) equilibrium outcomes with communication which Pareto dominate all (symmetric) equilibrium outcomes without communication. Players' privately known types change over time. This class of models has a natural application to oligopoly markets in which firms privately know their production possibilities (as in the papers mentioned above), or repeated auctions in which buyers' values change across time (as in Skrzypacz and Hopenhayn (2004)). In these contexts, whether communication helps or not has important practical and legal implications.

The role of communication here is intuitive. Players are able to aggregate private information through communication and thereby, to coordinate their

[1]These models study discounted games and so differ from the extensive literature on undiscounted repeated games of incomplete information. See Aumann and Maschler (1995) for a survey.

behavior. However, it does not directly follow that communication *actually* helps, as players always have an incentive to lie about their private information. Despite this, I show that there exists an equilibrium in which players communicate honestly most of the time and hence improve the equilibrium payoffs.

The stage game is a prisoner's dilemma with incomplete information. The payoff to a cooperating player when the other defects takes one of the two values: high or low. The pair of actions that maximizes the sum of payoffs thus depends on the payoff "shocks". If the negative payoff for one player is low, then the sum of payoffs is maximized by that player cooperating and the other defecting, rather than both players cooperating. I compare two cases: one in which communication is possible and the other in which it is not. In the case where communication is not possible, each individual decides her action after observing her own private payoff realization. In the other case, players strategically exchange (cheap talk) messages in each period after the shocks are realized and before the stage game actions are chosen. The messages are public and observed without any errors.

The main result is that there exists a nearly efficient payoff vector that is achieved as an equilibrium outcome when communication is possible and players are patient, but cannot be achieved without communication no matter how patient the players are. The second part is a simple corollary of a result that the feasible set without communication is a proper subset of the feasible set with communication. The first part is shown by construction. The equilibrium is described by a two-state automaton. There are two states: regular and punishment. In the regular state, players report their payoffs truthfully, and they choose the actions to maximize the sum of payoffs, given the reports. Any deviation on actions results in the state shift. Reporting that one gets high payoff shock also triggers the state change, with a small but strictly positive probability. In the punishment state, players play the stage game Nash equilibrium, *i.e.* mutual defections. The punishment state is the unique absorbing state.

Following Athey and Bagwell (2001), there has been a growing body of literature on discounted repeated games with incomplete information, mostly in specific models of oligopolies. This includes Athey and Bagwell (2001, 2008), Athey, Bagwell and Sanchirico (2004), Hörner and Jamison (2007) and Olszewski and Safronov (2012).

The paper is organized as follows. I set up the model in Section 2. The main result is stated in Section 3. The result is proved in the next two sections. Section 4 provide considers the case where communication is not possible and establishes an upper bound to the set of attainable payoffs. This bound is independent of the discount rate and is below the full information first-best payoff. Section 5 is devoted to the case where communication is possible and shows that when players are patient, the full information benchmark payoff can be approximated. Section 6 concludes.

2 Preliminaries

2.1 The Stage Game

In this paper, I consider the following prisoners' dilemma with incomplete information, denoted by G:

	C	D
C	$1, 1$	$-\theta_1, 1+g$
D	$1+g, -\theta_2$	$0, 0$

Figure 1: The payoff matrix

where θ_i is a *payoff shock* that is privately observed by player i only. Each θ_i can take on two values: l and h.

It is assumed that $g > 0$, $h > l > 0$ and that

$$1 + g - h < 2 < 1 + g - l \tag{1}$$

Notice that if the realized $\theta = (\theta_1, \theta_2) = (h, h)$, then the action pair (C, C) maximizes the sum of payoffs. If the realized $\theta = (l, h)$ (resp. (h, l)), then (C, D) (resp. (D, C)) maximizes the sum of payoffs. Finally, if the realized $\theta = (l, l)$, then either (D, C) or (C, D) maximizes the sum.

I note that the game described above is one of private values—player i's payoff depends only on θ_i and not θ_{-i}. (Throughout, the symbol $-i$ denotes player $3 - i$.)

Payoff shocks are drawn from a joint distribution $q \in \Delta$ where Δ denotes the set of distributions over $\Theta = \Theta_1 \times \Theta_2 = \{l, h\} \times \{l, h\}$. Notice that correlation among types is allowed. I assume that the distribution q has full support. Moreover, assume that for each i,

$$1 + g - [q_i(h)h + q_i(l)l] < 2 \tag{2}$$

and

$$l < \frac{q_i(h)}{q_i(l)} < q_i(h)h + q_i(l)l \tag{3}$$

where $q_i(\theta_i)$ denotes the marginal distribution that the payoff shock of player i is $\theta_i \in \{l, h\}$. I do not assume that q is commonly known.

Condition (2) implies that *ex ante* the game looks like a usual prisoner's dilemma, in the sense that (C, C) is more efficient than alternating between (C, D) and (D, C). Condition (3) is satisfied whenever, given the other parameters, l is small and h is big enough.

Let $A_i = \{C, D\}$ and $A = A_1 \times A_2$. A pure strategy of player i in G is a function

$$s_i : \Theta_i \to A_i$$

Let S_i denote the set of pure strategies of player i and let $S = S_1 \times S_2$. Then $u_i[q] : S \to \mathbb{R}$ is the *expected* payoff for player i, computed using the distribution $q \in \Delta$.

Define the set of feasible payoffs $\mathcal{F}[q] \subset \mathbb{R}^2$ by

$$\mathcal{F}[q] = \text{co}\{u[q](s) \mid s \in S\} \tag{4}$$

where $u[q] = (u_1[q], u_2[q])$ and co denotes the convex hull of a set.

2.2 Dynamic Game

Consider an infinite horizon dynamic game version $G(\delta)$ of the stage game in which both players discount future payoffs using a discount factor $\delta \in (0, 1)$. I consider a situation in which the distribution of player types may evolve over time.

Assume the distribution of $\theta \in \Theta$ at period $t+1$ is determined according to the history of θ realizations up to period t. Formally, let $q^{t+1}(\theta^{t+1} \mid \boldsymbol{\theta}^t)$ be the distribution at period $t+1$ given the history $\boldsymbol{\theta}^t = (\theta^1, \theta^2, ...\theta^t)$ of realizations up to period t.[2]

Assume that players are anonymous in the following sense.

Assumption 1. For any $\boldsymbol{\theta}^t, \boldsymbol{\theta'}^t \in \boldsymbol{\Theta}^t$ that satisfy $\boldsymbol{\theta}_i^t = \boldsymbol{\theta'}_{-i}^t$ and $\boldsymbol{\theta}_{-i}^t = \boldsymbol{\theta'}_i^t$,

$$q_i^{t+1}(\cdot \mid \boldsymbol{\theta}^t) = q_{-i}^{t+1}(\cdot \mid \boldsymbol{\theta'}^t)$$

and

$$q_{-i}^{t+1}(\cdot \mid \boldsymbol{\theta}^t) = q_i^{t+1}(\cdot \mid \boldsymbol{\theta'}^t)$$

The specification permits (i) processes that are independent and identical across periods (henceforth called *i.i.d.* processes) where for all t and histories $\boldsymbol{\theta}^t$, $q^{t+1}(\theta^{t+1} \mid \boldsymbol{\theta}^t) = q$; (ii) Markov processes where $q^{t+1}(\theta^{t+1} \mid \boldsymbol{\theta}^t) = q(\theta^{t+1} \mid$

[2] Hereafter, subscripts indicate players, superscripts indicate periods, and bold letters indicate histories.

θ^t). But it also permits more general processes like AR-2 processes where $q^{t+1}(\theta^{t+1} \mid \boldsymbol{\theta}^t) = q(\theta^{t+1} \mid \theta^t, \theta^{t-1})$ as well as many others.

Let

$$Q = \text{cl}\{q \in \Delta(\Theta) \mid \exists(t, \boldsymbol{\theta}^t) \; s.t. \; q(\cdot) = q^{t+1}(\cdot|\boldsymbol{\theta}^t)\}$$

where cl denotes the closure of a set. I make the following assumptions about Q.

Assumption 2.

$$Q \subset \text{Int}\Delta$$

where Int denotes the interior of a set.

This assumption implies that any $q \in Q$ has full support. Many stochastic processes satisfy the assumption. The simplest example is an *i.i.d.* process. Suppose that every period the shocks are drawn from a distribution $q \in \text{Int}\Delta$. In this case, $Q = \{q\} \subset \text{Int}\Delta$. A globally stable Markov process whose initial distribution has full support satisfies the assumption as well.[3]

A second assumption is the following.

Assumption 3. Any $q \in Q$ satisfies (2) and (3).

Finally, I assume that the distribution is sufficiently close to an *i.i.d.* process that is also *independent* across players (hereafter, referred as a *doubly i.i.d.* process). Let $\Delta_{Ind} \subset \Delta$ be the subset of distributions over Θ that are independent, that is, every $p \in \Delta_{Ind}$ can be written as a product of two distributions over Θ_i. Analogously, let $Q_{Ind} \subset \Delta_{Ind}$ be the set of independent distributions that are interior, satisfy (2) and (3) and anonymous.

To specify how close a stochastic process in Q is to a distribution in Q_{Ind}, I use the Euclidian norm on Δ (since Θ has four elements, Δ can be viewed as the unit simplex in $\mathbb{R}^4$).

Definition 1. *The set $Q \subset \Delta$ is d-*close *to a doubly i.i.d. process if there exists a $p \in Q_{Ind}$ such that for any $q \in Q$,*

$$||q - p|| \leq d$$

Notice that the definition has two requirements. Every element of Q should be close to the *same* element of Q_{Ind}. Since a single element of $q \in \Delta$ generates a process that is *i.i.d.* over time—that is, for all t and $\boldsymbol{\theta}^t$, $q^{t+1}(\theta^{t+1} \mid \boldsymbol{\theta}^t) = q$—the first requirement is that Q is close to an *i.i.d.* process. The condition that $p \in Q_{Ind}$, then says that the *i.i.d.* process has the property that payoff shocks are also independent across players.

Of course, any *i.i.d.* process generated by $q \in Q_{Ind}$ itself satisfies the assumptions for any d. Other examples that satisfy these assumptions are provided below.

[3]On the other hand, there exist irreducible processes that do not satisfy the assumption.

Example 1. Let $g = 2$, $h = 5/2$ and $l = 1/2$. Let a distribution p be $p_i(\theta_i) = 1/2$ for any i and θ_i. Notice $p \in Q_{Ind}$.

Let $d' > 0$ be a small number. Consider a Markov process that evolves according to

$$q(\theta^{t+1}|\theta^t) = \begin{cases} 1/4 + 3d' & \text{if } \theta^{t+1} = \theta^t \\ 1/4 - d' & \text{if } \theta^{t+1} \neq \theta^t \end{cases}$$

with the initial distribution p. Then, $Q = \{p\} \bigcup \{\cup_{\theta \in \Theta} q(\cdot|\theta)\}$. It is clear that Q satisfies Assumption 1, 2 and 3.

Finally, notice

$$\max_{q \in Q} ||q - p|| = 2\sqrt{3}d'$$

Hence, for any $d \geq 2\sqrt{3}d'$, the process is d-close to p.

Example 2. Let g, h, l and p be the same as the above example. Let $d' > 0$ be a small number. Consider an AR-2 process that evolves according to

$$q(\theta^{t+1}|\theta^t, \theta^{t-1}) = \begin{cases} 1/4 + 3d' & \text{if } \theta^{t+1} = \theta^t = \theta^{t-1} \\ 1/4 - d' & \text{if } \theta^{t+1} \neq \theta^t \text{ and } \theta^t = \theta^{t-1} \\ 1/4 & \text{if } \theta^t \neq \theta^{t-1} \end{cases}$$

with the initial distributions at $t = 1$ and $t = 2$ being $q(\cdot|\theta^0, \theta^{-1}) = p(\cdot) = q(\cdot|\theta^1, \theta^0)$.

For the same reasons as in the above example, the stochastic process satisfies all assumptions for $d \geq 2\sqrt{3}d'$.

The following is an example where there are multiple invariant distributions.

Example 3. Let g, h, l and p be the same as the above example. Let $d' > 0$ be a small number. Consider the following process:

$$q(\theta^{t+1}|\theta^1) = \begin{cases} 1/4 + 3d' & \text{if } \theta^{t+1} = \theta^1 \\ 1/4 - d' & \text{if } \theta^{t+1} \neq \theta^1 \end{cases}$$

with the initial distribution at $t = 1$, $q(\cdot|\theta^0) = p(\cdot)$.

For the same reasons as in the above example, the stochastic process satisfies all assumptions for $d \geq 2\sqrt{3}d'$.

Prior to choosing an action—C or D—in period t, player i observes the realization of his private payoff shock θ_i^t in that period. I assume that all past *actions* are observed by both players—there is *perfect monitoring of* *actions*—but that the payoff shocks remain private information. A *pure strategy* σ_i for player i in $G(\delta)$ is a sequence of function $\sigma_i^1, \sigma_i^2, ...$ where

$$\sigma_i^1 : \Theta_i \to A_i$$

and for $t > 1$,

$$\sigma_i^t : \boldsymbol{A}^{t-1} \times \boldsymbol{\Theta}_i^t \to \{C, D\}$$

where $\boldsymbol{A}^{t-1} = \times_{\tau=1}^{t-1} A$ is the history of actions and $\boldsymbol{\Theta}_i^t = \times_{\tau=1}^{t-1} \Theta_i$ is the history of i's payoff shocks.

Define the sequence of actions $a^t(\sigma)$ that result from a strategy pair $\sigma = (\sigma_1, \sigma_2)$ as follows: $a^1(\sigma, \theta) = (\sigma_1^1(\theta_1), \sigma_2^1(\theta_2))$ and in subsequent periods, $a^t(\sigma, \boldsymbol{\theta}^t) = \left(\sigma_1^t\left(\boldsymbol{a}^{t-1}, \boldsymbol{\theta}_1^t\right), \sigma_2^t\left(\boldsymbol{a}^{t-1}, \boldsymbol{\theta}_2^t\right)\right)$. Throughout, I allow players to use public randomization devices.

Given $\sigma = (\sigma_1, \sigma_2)$, let $V_i(\sigma)$ be the discounted average expected payoff of player i evaluated at the initial period. Hence,

$$V_i(\sigma) = (1-\delta) \sum_{t=0}^{\infty} \delta^t \mathbb{E}_{\boldsymbol{\theta}^t} U_i(a^t\left(\sigma, \theta^t\right), \theta_i)$$

where $U_i(\cdot, \theta_i)$ is the *ex post* utility function given the payoff shock θ_i and the expectation is taken over all histories $\boldsymbol{\theta}^t$. Let Σ_i be the set of pure strategies of player i and $\Sigma = \Sigma_1 \times \Sigma_2$.

Define $\mathcal{F}(\delta) \subset \mathbb{R}^2$ by

$$\mathcal{F}(\delta) = \text{co}\{V(\sigma) \mid \sigma \in \Sigma\} \tag{5}$$

where $V = (V_1, V_2)$. By anonymity, $\mathcal{F}(\delta)$ is symmetric.

2.3 Dynamic Game with Communication

Next consider an infinite horizon dynamic game version $G^*(\delta)$ of the stage game in which players can communicate. Specifically, they can send messages in a finite set M_i to each other in every period after learning their private payoff shocks but prior to choosing their actions. Future payoffs are again discounted using δ. Communication entails no cost so that it is "cheap talk". All messages are observed publicly and without any error.

The sequence of actions in each stage is as follows. First, players privately learn the realization of their payoff shocks θ_1 and θ_2. Then, they exchange messages simultaneously. In what follows, the message space is $M_i = \{L, H\}$ for both players. After messages are exchanged, players observe an outcome of a public randomization device.[4] Then they choose actions simultaneously.

[4]Throughout, I assume that players can utilize public randomization devices. This is an innocuous assumption because players can create public randomization devices through communication using "jointly controlled lotteries." See Aumann and Maschler (1995).

I again assume perfect monitoring of past actions and messages. In addition, each player observes his own current and past payoff shocks. A pure strategy in $G^* (\delta)$ is a pair (μ_i^*, σ_i^*) where $\mu_i^* = (\mu_i^{*1}, \mu_i^{*2}, ...)$ is a sequence of *message* strategies and $\sigma_i^* = (\sigma_i^{*1}, \sigma_i^{*2}, ...)$ is a sequence of *action* strategies. Formally,

$$\begin{aligned} \mu_i^{*1} &: \Theta_i \to M_i \\ \sigma_i^{*1} &: \Theta_i \times M \to A_i \end{aligned}$$

and for $t > 1$,

$$\begin{aligned} \mu_i^{*t} &: \boldsymbol{A}^{t-1} \times \boldsymbol{M}^{t-1} \times \boldsymbol{\Theta}_i^t \to M_i \\ \sigma_i^{*t} &: \boldsymbol{A}^{t-1} \times \boldsymbol{M}^{t} \times \boldsymbol{\Theta}_i^t \to A_i \end{aligned}$$

where $M = M_1 \times M_2$.

3 The Main Result

The main result of this paper is the following.

Theorem 1. *Suppose the stochastic process is almost doubly i.i.d. When players are sufficiently patient, there exist symmetric equilibrium payoffs of the game with communication, $G^* (\delta)$, which strictly Pareto dominate all symmetric equilibrium payoffs of the game without communication, $G (\delta)$. These payoffs of $G^* (\delta)$ cannot be approached by equilibrium payoffs of $G (\delta)$ as $\delta \to 1$.*

The result thus shows that to achieve certain "cooperative" payoffs, communication is *necessary.* Without communication, these payoffs cannot even be approached as $\delta \to 1$.

Theorem 1 is established as follows. First, consider a a one-shot game $\hat{G}$ in which the payoffs are the same as G but there is *full information*, that is, *both* payoff shocks are commonly known. A pure strategy for player i in $\hat{G}$ is of the form $\hat{s}_i : \Theta \to A_i$ and let $\hat{S}_i$ denote the set of pure strategies of player i in $\hat{G}$. Define $\hat{S} = \hat{S}_i \times \hat{S}_i$ and let $\hat{u}_i : \hat{S} \to \mathbb{R}$ be the expected stage game payoff for player i evaluated by a distribution $q \in \Delta$ and let $\hat{u} = (\hat{u}_1, \hat{u}_2)$. Then define

$$\hat{\mathcal{F}}[q] = \text{co}\{\hat{u}(\hat{s}) \mid \hat{s} \in \hat{S}\} \tag{6}$$

Now, for a set $F \subset \mathbb{R}^2$, let

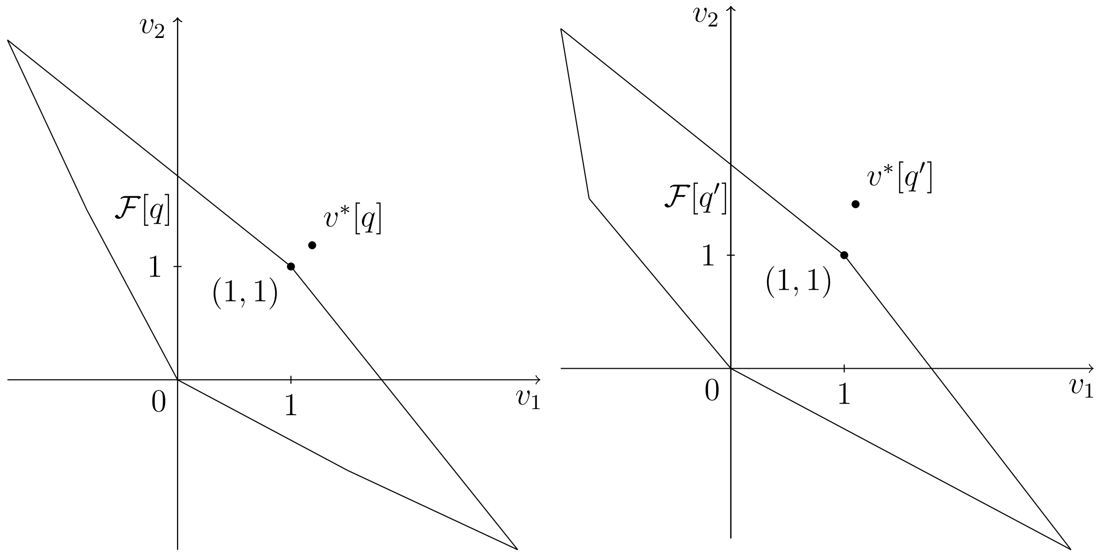
$$\overline{w}(F) = \sup \{v_1 + v_2 : (v_1, v_2) \in F\}$$

Then, $\overline{w}\left(\hat{\mathcal{F}}[q]\right)$ is the "first-best" benchmark payoff of $\hat{G}$ under a distribution $q \in \Delta$ since it is derived assuming that both players' information is common. Let $v^*[q] = (v_1^*[q], v_2^*[q]) \in \hat{\mathcal{F}}[q]$ satisfy $\overline{w}\left(\hat{\mathcal{F}}[q]\right) = v_1^*[q] + v_2^*[q]$.

Consider the dynamic version of full information game $\hat{G}(\delta)$ and let

$$W^*(\delta) = (1-\delta)\sum_{t=0}^{\infty} \delta^t \mathbb{E}_{\boldsymbol{\theta}^t} \overline{w}\left(\hat{\mathcal{F}}\left[q^t\left(\theta^t|\boldsymbol{\theta}^{t-1}\right)\right]\right)$$

where the expectation is taken over all histories $\boldsymbol{\theta}^t$.



The formal proof of the theorem then results from the following two propositions, whose proofs appear in the next two sections. The first proposition says that if there is no communication, the first-best benchmark cannot be approached, as $\delta \to 1$.

Proposition 1. *There exists an $\underline{\varepsilon} > 0$ such that for any $\delta \in (0,1)$,*

$$\overline{w}\left(\mathcal{F}(\delta)\right) \leq W^*(\delta) - \underline{\varepsilon}$$

where $\mathcal{F}(\delta)$ is given by (5).

The second proposition says that if there is communication, then the first-best benchmark can be approached, as the discount factor $\delta \to 1$ and the distance to a doubly *i.i.d.* process $d \to 0$.

Proposition 2. *For any $\varepsilon > 0$, there exists a $\bar{d} > 0$ and $\underline{\delta} \in (0,1)$ such that for any $\delta \in (\underline{\delta}, 1)$ and $d \in [0, \bar{d})$, for any game with communication, $G^*(\delta)$, with a stochastic process which is d-close to being doubly i.i.d., there exists an equilibrium whose payoffs sum to greater than $W^*(\delta) - \varepsilon$.*

4 Without Communication

In this section, I prove Proposition 1. Towards this purpose, first I show that for any distribution $q \in \text{Int}\Delta$,

$$\overline{w}\left(\mathcal{F}[q]\right) < \overline{w}\left(\hat{\mathcal{F}}[q]\right) \tag{7}$$

where $\mathcal{F}[q]$ and $\hat{\mathcal{F}}[q]$ are given by (4) and (6) respectively.

To see this, notice that in $\hat{G}$, if realized $\theta = (h, h)$, then the action pair (C, C) maximizes the sum of payoffs while if realized $\theta = (h, l)$, then (D, C) maximizes the sum of payoffs. However, in game G, a player's pure strategy depends only on his own payoff shock. Since for any set set $F \subset \mathbb{R}^2$, $\overline{w}(F) = \overline{w}(\text{co}F)$, and hence the maximum of payoffs that is feasible by pure strategy in G coincides to $\bar{w}(\mathcal{F}(q))$. Thus, if both (h, h) and (h, l) occur with positive probability, then (7) holds.

For each $q \in \Delta$, define

$$\eta[q] = \overline{w}\left(\hat{\mathcal{F}}[q]\right) - \overline{w}\left(\mathcal{F}[q]\right)$$

Notice, if q has full support, then $\eta[q] > 0$. Let $\underline{\eta} = \min_{q \in Q} \eta[q]$. Since Q is a compact subset of $\text{Int}\Delta$, $\underline{\eta}$ exists and is strictly positive. Now consider a dynamic game $\check{G}(\delta)$ in which realizations are still private information but the distribution in every period is commonly known, and let

$$\check{W}(\delta) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t \mathbb{E}_{\boldsymbol{\theta}^t} \overline{w}\left(\mathcal{F}\left[q^t\left(\theta^t | \boldsymbol{\theta}^{t-1}\right)\right]\right)$$

Since every feasible strategy in $G(\delta)$ is also feasible in $\check{G}(\delta)$,

$$\check{W}(\delta) \geq \overline{w}\left(\mathcal{F}(\delta)\right)$$

Now, notice that

$$\begin{aligned}
& W^*(\delta) - \overline{w}\left(\mathcal{F}(\delta)\right) \\
\geq\ & W^*(\delta) - \check{W}(\delta) \\
=\ & (1 - \delta) \sum_{t=0}^{\infty} \delta^t \mathbb{E}_{\boldsymbol{\theta}^t} \left[\overline{w}\left(\hat{\mathcal{F}}\left[q^t\left(\theta^t | \boldsymbol{\theta}^{t-1}\right)\right]\right) - \overline{w}\left(\mathcal{F}\left[q^t\left(\theta^t | \boldsymbol{\theta}^{t-1}\right)\right]\right)\right] \\
\geq\ & (1 - \delta) \sum_{t=0}^{\infty} \delta^t \underline{\eta} \\
=\ & \underline{\eta}
\end{aligned}$$

Let $\underline{\varepsilon} = \underline{\eta}$, and the claim is shown.

5 With Communication

5.1 Equilibrium Strategies

In this section, I prove Proposition 2. The proof is by construction. First I propose a candidate equilibrium whose sum of average payoffs are arbitrary close to $W^*(\delta)$. Then I show the candidate equilibrium actually constitutes an equilibrium for sufficiently large δ and sufficiently small d.

The idea of construction is closely related to "price war" equilibrium of Athey, Bagwell and Sanchirico (2004). The equilibrium strategies are described by two-state automata. One state is referred as regular and the other as punishment. The strategies in each state and the transition between states are described in the following.

The Regular State

In this state, message of each player only depends on his payoff shock, and actions depend on current messages and the outcome of a public randomization device. In the messaging stage, both players are *truthful*, that is, $\mu_i^*(h) = H$ and $\mu_i^*(l) = L$ and so each reports his own payoff shock truthfully.

After messages are exchanged, players jointly observe a uniformly distributed random variable $\omega \in [0, 1]$. Players follow the *action rule* α^*, described as follows:

$$\alpha^*(m,\omega) = \begin{cases} (C,C) & \text{if } \ m = (H,H) \ \text{ and } \ \omega > \gamma \\ (D,D) & \text{if } \ m = (H,H) \ \text{ and } \ \omega \leq \gamma \\ (C,D) & \text{if } \ m = (L,H) \ \text{ or } \\ & \text{if } \ m = (L,L) \ \text{ and } \ \omega > 1/2 \\ (D,C) & \text{if } \ m = (H,L) \ \text{ or } \\ & \text{if } \ m = (L,L) \ \text{ and } \ \omega \leq 1/2 \end{cases}$$

where $\gamma > 0$ will be specified later. Notice the sum of the stage game payoffs is close to the first-best benchmark when γ is small enough.

The Punishment State

In the punishment state, neither of payoff shocks nor messages affect actions. Players just play the unique static Nash equilibrium, which is, (D, D).

Transitions

In the initial period, the state is regular. The state remain unchanged until one of the following event happens. (1) A deviation of stage game is observed.

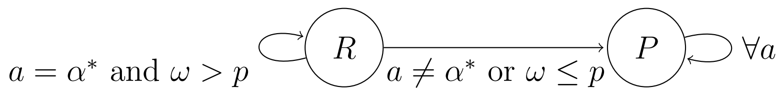


Figure 2: Transition when reports are (H, H)

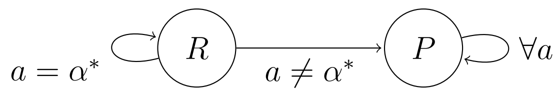


Figure 3: Transition when reports are not (H, H)

(2) If the pair of messages is (H, H) and $\omega \leq \gamma$. If either (1) or (2) occurs, the state shifts to the punishment state.

The transition is illustrated in the following two figures. Here, $a \in A$ is an action pair that is actually played.

Lifetime Utility

I show that as γ goes to zero, the sum of payoffs of the candidate equilibrium converges to the first-best benchmark payoff $W^*(\delta)$. To see this, just notice that the probability that the state stays in the regular state at period t is at least $(1 - \gamma\bar{q}(h, h)^2)^t$ where $\bar{q} \in Q$ is the distribution whose probability of being $(\theta_1, \theta_2) = (h, h)$ is the highest among Q. Clearly,

$$\lim_{\gamma \to 0}(1 - \gamma\bar{q}(h, h)^2)^t = 0$$

Let $W(\gamma, \delta)$ be the sum of the expected payoffs when both players follow the candidate equilibrium. Then,

$$W^*(\delta) - W(\gamma, \delta) \leq (1 - \delta)\sum_{t=0}^{\infty}\delta^t[1 - (1 - \gamma\bar{q}(h, h)^2)^t](1 + g - l) \to 0$$

as $\gamma \to 0$.

5.2 Optimality of Stage Game Actions

Here, I show that the prescribed strategies constitute an equilibrium for large enough δ, small enough d and small enough γ. In the punishment state, clearly there is no profitable deviation. In the regular state, there are two kinds of deviations: one concerns actions and the other messages. First, I assume that both send truthful messages and the other player follows the action rule α^*, and show that for appropriately chosen (δ, γ), there is no profitable deviation from the action rule α^*.

Lemma 1. *Suppose both players are truthful and that player $-i$ follows α^*. Then, there exists a $(\bar{d}, \overline{\gamma}, \underline{\delta}) \in (0, \infty) \times (0, 1] \times [0, 1)$ such that for any (d, γ, δ) that satisfies $d < \bar{d}$, $\gamma < \overline{\gamma}$ and $\delta > \underline{\delta}$, following α^* is optimal for player i.*

Proof. Let the lifetime utility when both players follow the rule α^* denoted by V^* (for notational ease, I am suppressing the dependence of the lifetime utility payoff on t and $\boldsymbol{\theta}^{t-1}$). Notice that in each stage game, a player can get at most $1 + g$ while at least $-h$. Then, the payoff of a deviation is at most $(1-\delta)(1+g)$, while the payoff of not deviating is at least $(1-\delta)(-h) + \delta V^*$.

Now, to show that the deviation is not profitable, it is sufficient to show

$$-(1-\delta)(1+g-h) + \delta V^* > 0$$

This inequality is satisfied if δ is sufficiently large and $V^* > 0$. To see that $V^* > 0$, let the lifetime utility when stochastic process follows a doubly *i.i.d.* process p be $V^*[p]$ and show that $V^*[p] > 0$ for sufficiently large δ. It is clear that

$$\lim_{d \to 0} V^* = V^*[p] \tag{8}$$

and the claim follows if I show $V^*[p] > 0$. Now notice

$$\begin{aligned} V^*[p] &= (1-\delta)(v^*[p] - \gamma p_h^2) + \delta[p_l^2 + 2p_h p_l + (1-\gamma)p_h^2]V^*[p] \\ &= (1-\delta)(v^*[p] - \gamma p_h^2) + \delta[1 - \gamma p_h^2]V^*[p] \end{aligned}$$

where $v^*[p]$ is the best symmetric payoff given distribution p. Notice $v^*[p] > 0$. Thus

$$V^*[p] = \frac{1-\delta}{1-\delta(1-\gamma p_h^2)}(v^*[p] - \gamma p_h^2) > 0 \tag{9}$$

for sufficiently small p. $\square$

5.3 Optimality of Messages

Here I show that for appropriately large δ and small d, for any $\gamma > 0$, it is optimal for player i to tell the truth if his opponent is also doing so and both players play according to α^*.

Let $V^*(m_i, \theta_i)$ be the *interim* lifetime utility of player i when player i gets a payoff shock θ_i and sends a message m_i (again I am suppressing the dependence of the interim payoff on t and $\boldsymbol{\theta}^{t-1}$). Then, I have

$$V^*(H, \theta_i) = (1-\gamma)q_{-i}(h)(1-\delta+\delta\mathbb{E}V^*) + q_{-i}(l)[(1-\delta)(1+g) + \delta\mathbb{E}V^*]$$

$$V^*(L,\theta_i) = q_{-i}(h)[(1-\delta)(-\theta_i)+\delta\mathbb{E}V^*]+q_{-i}(l)\left((1-\delta)\frac{1+g-\theta_i}{2}+\delta\mathbb{E}V^*\right)$$

where $\mathbb{E}V^*$ denotes the expected value of future payoffs and the expectation is taken over all future realizations of the payoff shocks. Then, truth-telling is optimal if

$$V^*(L,l) > V^*(L,h)$$

and

$$V^*(H,h) > V^*(L,h)$$

hold.

The claim is established by the following two steps. First, I show that truth-telling is optimal if the stochastic process is the doubly *i.i.d.* process $\{p\}$. Then, in the second step, I show that if Q is close to p, the payoff for these processes are also close for any strategies. Combining these, it is shown that truth-telling is optimal for the original process Q.

Lemma 2. *Suppose the stochastic process is given by a doubly i.i.d. process* p*. Suppose both players follow the action rule* α^* *and that player* $-i$ *is truthful. Then, for any* $\gamma \in (0,1)$*, there exists a* $\underline{\delta} \in [0,1)$ *such that for any* $\delta > \underline{\delta}$*, truth-telling is optimal for player* i*.*

The idea of the proof is self-selection. In terms of stage game payoffs, sending message H dominates L, regardless of realizations. Hence, to assure truth-telling, there must be a punishment for sending message H. The punishment must be strong enough to prevent L type to send message H, but at the same time, weak enough for H type to send message H. By choosing γ and δ appropriately, I can find the punishment.

Proof. For a doubly *i.i.d.* process p, let $V_i^*[p](m_i,\theta_i)$ be the *interim* payoff of player i who gets a payoff shock $\theta_i \in \{l,h\}$ and sends a message $m_i \in \{L,H\}$. Then,

$$V_i^*[p](H,\theta_i) = (1-\gamma)p_h[1-\delta+\delta V^*[p]] + p_l[(1-\delta)(1+g)+\delta V^*[p]] \quad (10)$$

$$V_i^*[p](L,\theta_i) = p_h[(1-\delta)(-\theta_i)+\delta V^*[p]] + p_l\left[(1-\delta)\frac{1+g-\theta_i}{2}+\delta V^*[p]\right] \quad (11)$$

where $p_h = p_{-i}(h)$, $p_l = p_{-i}(l)$ and $V^*[p]$ is given by (9). By anonymity, $p_h = p_i(h)$ and $p_l = p_i(l)$. Now, truth-telling is optimal if

$$V_i^*[p](L,h) < V_i^*[p](H,h) = V_i^*[p](H,l) < V_i^*[p](L,l)$$

is satisfied. From (10) and (11),

$$V_i^*[p](H,\theta_i) - V_i^*[p](L,\theta_i) = (1-\delta)\left(y(\theta_i)-\gamma p_h\right) - \delta\gamma p_h V^*[p] \quad (12)$$

where

$$y(\theta_i) = 1 + p_l g + p_h \theta_i - p_l \frac{1 + g - \theta_i}{2}$$

Combining these, truth-telling is optimal if

$$(1-\delta)[y(l) + \gamma p_h] < \delta \gamma p_h V^*[p] < (1-\delta)[y(h) + \gamma p_h] \tag{13}$$

holds.

To see (13) actually holds, first let us show that

$$0 < y(l) < y(h)$$

To prove the first inequality, see

$$\begin{aligned} y(l) &= 1 - \frac{p_l}{2} + \frac{p_l}{2} g + \left(1 - \frac{p_l}{2}\right) l \\ &> 1 - \frac{p_l}{2} + \frac{p_l}{2}(1 + l) + \left(1 - \frac{p_l}{2}\right) l \\ &= 1 + l \\ &> 0 \end{aligned}$$

where the first inequality follows from (2). The second inequality follows from

$$y(h) - y(l) = \left(p_h + \frac{p_l}{2}\right)(h - l) > 0$$

Substituting (9) into (13) gives us

$$y(l) < \gamma p_h + \frac{\delta \gamma p_h^2}{1 - \delta(1 - \gamma p_h^2)} (v^*[p] - \gamma p_h^2) < y(h)$$

Then, by doing some algebra, it is shown that (13) is equivalent to

$$\frac{y(l) - \gamma p_h}{(1 - \gamma p_h^2) y(l) - \gamma p_h + \gamma p_h v^*[p]} < \delta < \frac{y(h) - \gamma p_h}{(1 - \gamma p_h^2) y(h) - \gamma p_h + \gamma p_h v^*[p]} \tag{14}$$

Now, suppose

$$\frac{y(l) - \gamma p_h}{(1 - \gamma p_h^2) y(l) - \gamma p_h + \gamma p_h v^*[p]} < 1 < \frac{y(h) - \gamma p_h}{(1 - \gamma p_h^2) y(h) - \gamma p_h + \gamma p_h v^*[p]} \tag{15}$$

holds. Then, let

$$\underline{\delta} = \frac{y(l) - \gamma p_h}{(1 - \gamma p_h^2) y(l) - \gamma p_h + \gamma p_h v^*[p]}$$

and for any $\delta \in (\underline{\delta}, 1)$, (14) holds. This shows that truth-telling is an optimal response.

To complete the proof, I show that assumptions (2) and (3) imply (13). The first inequality is shown as follows. Notice

$$y(l) - (1 - \gamma p_h^2)y(l) - \gamma p_h v^*[p] = \gamma p_h (p_h y(l) - v^*[p])$$

Now since $\gamma > 0$, the (15) holds if $v^*[p] - p_h y(l) > 0$. This is shown as follows:

$$\begin{aligned}
& v^* - p_h y(l) \\
= \;& (1+g-l)p_h p_l + p_l^2 \frac{1+g-l}{2} - p_h p_l (1+g) - p_h^2 l + p_h p_l \frac{1+g-l}{2} \\
= \;& -p_h l + p_l \frac{1+g-l}{2} \\
> \;& -p_h l + p_l \\
> \;& 0
\end{aligned}$$

where the second inequality follows from (3). The second inequality results from

$$y(h) - (1 - \gamma p_h^2)y(h) - \gamma p_h v^*[p] = \gamma p_h (p_h y(h) - v^*[p])$$

and

$$\begin{aligned}
& v^*[p] - p_h y(h) \\
= \;& p_h (p_h h + p_l l) - p_l \frac{1+g-p_h h - p_l l}{2} \\
> \;& p_h (p_h h + p_l l) - p_l \\
> \;& 0
\end{aligned}$$

where the first and second inequality above are the result of (2) and (3) respectively. □

The second step is to show that the payoff when the process evolves according to a process q that is sufficiently close to the *i.i.d.* process p is sufficiently close to the payoff when the process evolves according to the *i.i.d.* process. Then, I have the following lemma.

Lemma 3. *For any δ and γ,*

$$\lim_{d \to 0} V^*(m_i, \theta_i) = V^*[p](m_i, \theta_i)$$

Proof. Notice that

$$\begin{aligned}
& V^*(H, \theta_i) - V^*[p](H, \theta_i) \\
= \;& (1-\gamma) q_{-i}(h)[1 - \delta + \delta \mathbb{E} V^*] + q_{-i}(l)[(1-\delta)(1+g) + \delta \mathbb{E} V^*] \\
& -(1-\gamma) p_h [1 - \delta + \delta V^*[p]] - p_l [(1-\delta)(1+g) + \delta V^*[p]] \\
= \;& (1-\delta)\left[(1-\gamma)(q_{-i}(h) - p_h) + (q_{-i}(l) - p_l)(1+g)\right] \\
& +\delta \left[(1-\gamma)(q_{-i}(h)\mathbb{E} V^* - p_h V^*[p]) + p_l [\mathbb{E} V^* - V^*[p]]\right]
\end{aligned}$$

The first term vanishes since

$$\lim_{d \to 0} q_{-i}(h) = p_h$$

The second term vanishes because of (8). The second equality is shown similarly. $\square$

Combining these two lemmas, I have the following result.

Lemma 4. *Suppose both players follow the action rule α^* and that player $-i$ is truthful. Then there exists a triple of $(\overline{d}, \underline{\delta}, \overline{\gamma}) \in (0, \infty) \times (0, 1) \times (0, 1)$ such that for any (δ, d, γ) that satisfies $d < \overline{d}$, $\delta > \underline{\delta}$, and $\gamma < \overline{\gamma}$, truth-telling is optimal for player i.*

This lemma, combined with Lemma 2 establishes Proposition 2.

6 Conclusion

This paper provides a simple model in which communication is *essential* to achieving better outcomes. There are a few possible extensions. First, the assumptions on parameter values made in (2) and (3) might be relaxed, perhaps by employing some other kinds of strategies. Review strategies (see Hörner and Jamison [11] and Escobar and Toikka [6]) and chip strategies (see Olszenski and Safronov [13]) are natural candidates. Second, I wish to study the *extent* to which communication helps, and under what conditions. For example, how does the extent of correlation among player types affect the value of communication? Finally, the role of communication has important policy implications in the context of oligopolies. I hope to study the role on communication in such contexts as well.

References

[1] Athey, S., and K. Bagwell (2001): "Optimal Collusion with Private Information", *RAND Journal of Economics*, 32, 428-465..

[2] Athey, S., and K. Bagwell (2008): "Collusion with Persistent Cost Shocks", *Econometrica*, 76, 493-540.

[3] Athey, S., K. Bagwell, and C. Sanchirico (2004): "Collusion and price rigidity," *Review of Economic Studies*, 71, 317-349.

[4] Aumann, R.J., and M. B. Maschler (1995): *Repeated Games with Incomplete Information*, The MIT Press.

[5] Compte, O, (1998): “Communication in Repeated Games with Imperfect Private Monitoring,” *Econometrica*, 66, 597-626.

[6] Escobar, J. F., and J. Toikka,(2012): “Efficiency in Games with Markovian Private Information”. mimeo.

[7] Fudenberg, D., D. Levine and E. Maskin (1994): “The Folk Theorem with Imperfect Public Monitoring,” *Econometrica*, 62, 997-1039.

[8] Fudenberg, D., and E. Maskin (1986): “The Folk Theorem in Repeated Games with Discounting or with Incomplete Information,” *Econometrica*, 54, 533-554.

[9] Fudenberg, D., and Y. Yamamoto (2010): “Repeated Games where the Payoffs and Monitoring Structure are Unknown”, *Econometrica*, 78, 1673-1710.

[10] Fudenberg, D., and Y. Yamamoto (2011): “Learning from Private Information in Noisy Repeated Games”, *Jornal of Economic Theory*, 146, 1733-1769.

[11] Hörner, J., and J. Jamison (2007): “Collusion with (Almost) No Information”, *RAND Journal of Economics*, 38, 804-822.

[12] Kandori, M., and M. Matsushima (1998): “Private Observation, Communication and Collusion,” *Econometrica*, 66, 627-652.

[13] Olszewski, W., and M. Safronov (2012): “Chip Strategies in Repeated Games with Incomplete Information”, mimeo.

[14] Skrzypacz, A., and H. Hopenhayn (2004): “Tacit Collusion in Repeated Auctions”, *Journal of Economic Theory*, 114, 153-169.

[15] Sugaya, T. (2011): “Folk Theorem in Repeated Games with Private Monitoring”, mimeo.

[16] Yamamoto, Y. (2012): “Individual Learning and Cooperation in Noisy Repeated Games”, mimeo.