Communication and Cooperation in Repeated Games*

Yu Awaya† and Vijay Krishna‡

October 1, 2018

Abstract

We study the role of communication in repeated games with private monitoring. We first show that without communication, the set of Nash equilibrium payoffs in such games is a subset of the set of ε-coarse correlated equilibrium payoffs (ε-CCE) of the underlying one-shot game. The value of ε depends on the discount factor and the quality of monitoring. We then identify conditions under which there are equilibria with "cheap talk" that result in nearly efficient payoffs outside the set ε-CCE. Thus, in our model, communication is necessary for cooperation.

1 Introduction

The proposition that communication is necessary for cooperation seems quite natural, even self-evident. Indeed, in the old testament story of the Tower of Babel, God thwarted the mortals' attempt to build a tower reaching the heavens merely by dividing the languages. The inability to communicate with each other was enough to doom mankind's building project. At a more earthly level, antitrust laws in many countries prohibit or restrict communication among firms. Again, the premise is that limiting communication limits collusion. Organizations seek to design internal communication protocols to improve performance. Many use subjective peer evaluations—nonverifiable communication among employees and managers—to compensate employees and it is felt that such schemes provide stronger incentives for hard work (cooperation).

*The research reported here was supported by a grant from the National Science Foundation (SES-1626783). We are grateful to Gabriel Carroll, Tim Roughgarden and Satoru Takahashi for their comments and suggestions. The referees and the editor also provided valuable input.

†Department of Economics, University of Rochester, E-mail: yuawaya@gmail.com.

‡Department of Economics, Penn State University, E-mail: vkrishna@psu.edu.

Despite its self-evident nature, it is not clear how one may formally establish the connection between communication and cooperation. One option is to consider the effects of *pre-play* communication in one-shot games. This allows players to coordinate and even correlate their play in the game.[1] But in many games of interest, this does not enlarge the set of equilibria to allow for cooperation. For instance, pre-play communication has no effect in the prisoners' dilemma. It also has no effect in a differentiated-product price-setting oligopoly with linear demands.[2]

In this paper we study the role of *in-play* communication in repeated games—the basic framework for analyzing the prospects for cooperation among self-interested parties. The main idea of the theory of repeated games is that players are willing to forgo short-term gains in order to reap future rewards. But this relies on the ability of players to monitor each other well. If monitoring is poor, cooperative outcomes are hard to sustain because players can cheat with impunity and we study whether, in these circumstances, communication can help. Specifically, we show how *in-play* communication improves the prospects for cooperation in repeated games with imperfect *private* monitoring. In such settings, players receive only noisy private signals about the actions of their rivals.[3]

Our main result (Theorem 6.1) identifies monitoring structures—the stochastic mapping between actions and signals—with the property that communication is necessary for cooperation.

Theorem *For any high but fixed discount factor, there exists a non-empty and open set of monitoring structures such that there is an equilibrium with communication whose welfare exceeds that from any equilibrium without communication.*

What kinds of monitoring structures lead to this conclusion? Two conditions are needed. First, private signals should be rather noisy so that in the absence of communication monitoring is poor. Second, private signals of players should be strongly correlated when they play efficient action profiles and less so otherwise.[4] The second condition is natural in many economic environments. In an earlier paper (Awaya and Krishna, 2016), we explored a price-setting duopoly in which firms' sales were correlated in this manner and showed how this arose naturally because of randomness in consumers' search costs. Another class of situations in which the second condition is natural is the following. Suppose that players can play cooperatively—expend effort, contribute to a public good or gather information—or play selfishly and free

[1]The set of equilibrium payoffs with pre-play communication is contained within with set of correlated equilibrium payoffs of the original game. This follows from Aumann (1987).

[2]This is a *potential game* in the sense of Monderer and Shapley (1996) and a result of Neyman's (1997) then implies that the set of correlated equilibria coincides with the unique Nash equilibrium.

[3]A classic example is Stigler's (1964) model of secret price cuts where firms choose prices that are not observed by other firms. The prices (actions) then stochastically determine firms' sales (signals). Each firm only observes its own sales and must infer its rivals' actions only via these.

[4]Example 2 below shows that communication does not help if the reverse is true—that is, if correlation actually increases following a deviation.

ride on other players. These choices result in noisy private signals about an unknown state of nature. The stochastic structure is such that when a player plays cooperatively, the signal is more informative about the state than if he or she were to play selfishly. In any such situation, the players' signals will be more correlated when all cooperate than when this is not the case.[5] This kind of structure is common to many economic situations of interest. For instance, Fleckinger (2012) and Deb, Li and Mukherjee (2016) study a moral hazard situation in which agents' outputs are highly correlated when all agents work and less so if anyone shirks. Deb et al. study the role of communication—in the form of peer evaluations—in providing incentives. Similarly, Gromb and Martimort (2007) study the incentives of, say, two experts to gather information. When both experts expend effort, they learn more about a true state of the world than otherwise. This also results in the features outlined above.

How exactly does communication facilitate cooperation in such settings? The basic idea is that players can monitor each other not only by what they "see"—the signals—but also by what they "hear"—the messages that are exchanged. But since the messages are just cheap talk—costless and unverifiable—one may wonder how these can be used for monitoring. We construct equilibria in which the messages are cross-checked to ensure truthful reporting of signals. Moreover, the cross-checking is sufficiently accurate so that deviating players find it difficult to lie effectively. These essential properties rely on the second condition on the monitoring structure mentioned above. We use these ideas to construct an equilibrium that is nearly efficient.

Our main result requires two steps. The first task is to find an effective bound on equilibrium payoffs that can be achieved without communication. But the model we study is that of a repeated game with private monitoring and there is no known characterization of the set of equilibrium payoffs. This is because with private monitoring each player knows only his own history (of past actions and signals) and has only noisy information about the private histories of other players. Since players' histories are not commonly known, these cannot be used as state variables in a recursive formulation of the equilibrium payoff set. In Section 4, we borrow an equilibrium notion from algorithmic game theory—that of a coarse correlated equilibrium—and are able to relate (in Proposition 4.1) the Nash equilibrium payoffs of the *repeated* game to the *coarse correlated* equilibrium payoffs of the *one-shot* game. The notion of a coarse correlated equilibrium was introduced by Moulin and Vial (1978).[6] The set of coarse correlated equilibria is larger than the set of correlated equilibria and so has less predictive power in one-shot games. But because it is very easy to compute, in some cases it is nevertheless useful in bounding the set of Nash equilibria—for instance, in congestion games (Roughgarden, 2016). Here we show that it is useful

[5] In Example 1 below, we provide a simple derivation of this structure.

[6] While a formal definition appears in Section 2, a coarse correlated equilibrium is a joint distribution over players' actions such that no player can gain by playing a pure action under the assumption that the other players will follow the marginal distribution over their actions.

in bounding the set of Nash equilibria of repeated games as well. Precisely,

Proposition *The set of Nash equilibrium payoffs of the repeated game without communication is a subset of the set of ε-coarse correlated equilibrium payoffs of the one-shot game.*

The "ε" is determined by the discount factor and the monitoring structure of the repeated game and we provide an explicit formula for this. When the monitoring quality is poor—it is hard for other players to detect a deviation—ε is small and the set of ε-coarse correlated equilibrium payoffs provides an effective bound to the set of equilibrium payoffs in the repeated game.

The second task is to show that with communication, equilibrium payoffs above the bound can be achieved. To show this, we construct a cooperative equilibrium explicitly in which, in every period, players publicly report the signals they have received (see Proposition 5.1). Players' reports are aggregated into a single "score" and the future course of play is completely determined by this summary statistic. Deviations result in low scores and trigger punishments with higher probability. The score function lets us construct an equilibrium with communication to be of a particularly simple "trigger-strategy" form. When signals are "very informative" about each other, the constructed equilibrium is nearly efficient.

We emphasize that the analysis in this paper is of a different nature than that underlying the so-called "folk theorems" (see Sugaya, 2015). These show that for a fixed monitoring structure, as players become increasingly patient, near-perfect collusion can be achieved in equilibrium. In this paper, we keep the discount factor fixed and change the monitoring structure so that the set of equilibria with communication is substantially larger than the set without. A difficulty here is that monitoring structures—which are stochastic mappings from actions to signals—are high-dimensional objects. We show, however, that only *two* easily computable parameters—one measuring how noisy the signals are and the other how strongly correlated they are—suffice to identify monitoring structures for which communication is necessary for cooperation.

Related literature

The current paper builds on our earlier work, Awaya and Krishna (2016), where we explored some of the same issues in the special context of Stigler's (1964) model of secret price cuts in a symmetric duopoly with noisy sales. This is, of course, the canonical example of a repeated game with private monitoring. Much of the analysis in our earlier paper, however, relied on the assumption that sales were (log-) normally distributed and that the two firms were symmetric. This paper considers *general* n-person finite games and *general* signal distributions. More important, the bound on payoffs without communication that is developed here is tighter than the bound constructed in the earlier paper as well as being easier to interpret.[7] Moreover, the construction of equilibria with communication is entirely different and does not rely on any symmetry among players. Finally, we show in this paper (see Example 2 below) that communication does not always help cooperation. Awaya (2014) explores some of the same issues in a repeated prisoners' dilemma in continuous time.

There is a vast literature on repeated games under different kinds of monitoring. Under *perfect* monitoring, given any fixed discount factor, the set of perfect equilibrium payoffs with and without communication is the same. Under *public* monitoring, again given any fixed discount factor, the set of (public) perfect equilibrium payoffs with and without communication is also the same. Thus, in these settings communication does not affect the set of equilibria.[8]

Compte (1998) and Kandori and Matsushima (1998) study repeated games with *private* monitoring allowing, as we do, for *communication* among players. In this setting, they show that the folk theorem holds—any individually rational and feasible outcome can be approximated as the discount factor tends to one. These results are derived under specific assumptions about the detectability of deviations by other players and can be satisfied only if there are at least three players. This line of research has been pursued by others as well (see Aoyagi (2002), Fudenberg and Levine (2007), Zheng (2008) and Obara (2009)) in environments different from, and sometimes more general than, those of Compte (1998) and Kandori and Matsushima (1998). Particularly related to the current paper is the work of Aoyagi (2002) and Zheng (2008) who assume, as we do, that players' signals exhibit greater correlation when efficient actions are played than when actions are inefficient. All of these papers thus show that communication is *sufficient* for cooperation when players are sufficiently patient. But as Kandori and Matsushima (1998) recognize, "One thing which we did *not* show is the *necessity* of communication for a folk theorem" (p. 648, their italics).

In a remarkable paper, Sugaya (2015) shows the surprising result that in very general environments, the folk theorem holds *without* communication. Thus, in fact, communication is *not* necessary for a folk theorem. Although Sugaya's result was preceded by folk theorems for some limiting cases where the monitoring was almost perfect or almost public, the generality of its scope was unanticipated.

Unlike the folk theorems, in our work we do not consider the limit of the set of equilibrium payoffs as players become arbitrarily patient. We study the set of equilibrium payoffs for a fixed discount factor. Key to our result is a method of bounding the set of payoffs without communication using the easily computable set of ε-coarse correlated equilibria. Pai, Roth and Ullman (2016) also develop a bound that depends on a measure of monitoring quality based on the computer-science notion of "differential privacy." But the bound so obtained applies to equilibrium payoffs with communication as well as those without, and so does not help in distinguishing between the two.

[7] See Section 4.4.

[8] To be precise, these equivalences require that in the absence of communication, players have access to a public randomizing device ("sunspots") and that all communication is public.

Sugaya and Wolitzky (2017) find sufficient conditions under which the equilibrium payoffs with private monitoring are bounded by the equilibrium payoffs with perfect monitoring. This bound again applies whether or not there is communication and so is also unable to distinguish between the two.

Spector (2015) shows that communication can be beneficial in a model of price competition with private monitoring. Firms see their own current sales but, unlike in our model, can see other firms' sales with some delay. Communication is helpful in reducing this delay in monitoring. In our model, all communication is pure cheap talk—private signals remain so forever.

Kandori (2003) establishes a folk theorem with communication in a repeated game with public monitoring. With public monitoring, the only useful communication concerns the privately known actions that players have taken. In equilibrium such actions remain private when players randomize and the outcomes of these randomizations are known only to the player in question. Kandori shows how with private strategies and communication, a folk theorem may be established under weaker conditions than required when only public strategies are used. But again, this does not show the necessity of communication since it is not known what can be achieved with private strategies without communication. Rahman (2014) derives a similar result in a duopoly model.

The role of communication in fostering cooperation has also been the subject of numerous experiments in varied informational settings. Of particular interest is the work of Ostrom et al. (1994, Chapter 7) who find that in-play communication in repeated common-pool resource games leads to greater cooperation than does preplay communication. In recent work, Aryal, Ciliberto and Leyden (2018) study how public communication affects collusion among airlines.

The remainder of the paper is organized as follows. The next section outlines the formal model of repeated games with private monitoring. To motivate the subsequent analysis, in Section 3 we present some of the main ideas, as well as some subtleties, by means of some simple examples. Section 4 analyzes the repeated game without communication whereas Section 5 does the same with communication. The benefits of communication are established in Section 6. Section 7 concludes. Appendix A contains omitted proofs from Section 4 regarding the no-communication bound. Appendix B analyzes the score function that forms the basis of the equilibrium with communication that is constructed in Section 5.

2 Preliminaries

As mentioned in the introduction, we study repeated games with private monitoring.

Stage game The underlying game is defined by $\left(I, (A_i, Y_i, w_i)_{i\in I}, q\right)$ where $I = \{1, 2, ..., n\}$ is the set of players, A_i is a finite set of *actions* available to player i

and Y_i is a finite set of *signals* that i may observe. The actions of all the players $a \equiv (a_1, a_2, ..., a_n) \in A \equiv \times_i A_i$ together determine $q(\cdot \mid a) \in \Delta(Y)$, a probability distribution over the signals of all players.[9] A vector of signals $y \in Y$ is drawn from this distribution and player i only observes y_i. Player i's payoff is then given by the function $w_i : A_i \times Y_i \to \mathbb{R}$ so that i's payoff depends on other players' actions only via the induced signal distribution $q(\cdot \mid a)$. We will refer to $w_i(a_i, y_i)$ as i's *ex post* payoff.[10]

Prior to any signal realizations, the *expected* payoff of player i is then given by the function $u_i : A \to \mathbb{R}$, defined by

$$u_i(a) = \sum_{y_i \in Y_i} w_i(a_i, y_i)\, q_i(y_i \mid a)$$

where $q_i(\cdot \mid a) \in \Delta(Y_i)$ is the marginal distribution of $q(\cdot \mid a)$ on Y_i so that $q_i(y_i \mid a) = \sum_{y_{-i} \in Y_{-i}} q(y_i, y_{-i} \mid a)$. As usual, $\|u\|_\infty$ denotes the sup-norm of u. In what follows, we will merely specify the expected payoff functions u_i not the underlying ex post payoff functions w_i. The latter can be derived from the former for generic signal distributions—specifically, as long as $\{q_i(\cdot \mid a) : a \in A\}$ is a linearly independent set of vectors.

We refer to $G \equiv (A_i, u_i)_{i \in I}$ as the *stage game.* The set of *feasible* payoffs in G is $\mathcal{F} = \operatorname{co} u(A)$, the convex hull of the range of u. A payoff vector $v^* \in \mathcal{F}$ is (strongly) *efficient* if there does not exist a feasible $v \neq v^*$ such that $v \geq v^*$.

The collection $\{q(\cdot \mid a)\}_{a \in A}$ is referred to as the *monitoring structure.* We suppose throughout that for all i, the marginal distribution over i's signals, $q_i(\cdot \mid a) \in \Delta(Y_i)$ has *full support*, that is, for all $y_i \in Y_i$ and $a \in A$,

$$q_i(y_i \mid a) > 0 \tag{1}$$

Quality of monitoring Let $q_{-i}(\cdot \mid a) \in \Delta(Y_{-i})$ be the marginal distribution of $q(\cdot \mid a) \in \Delta(Y)$ over the joint signals of the players $j \neq i$. The *quality* of a monitoring structure q is defined as

$$\eta = \max_i \max_{a, a_i'} \|q_{-i}(\cdot \mid a) - q_{-i}(\cdot \mid a_i', a_{-i})\|_{TV} \tag{2}$$

[9] We adopt the following notational conventions throughout: capital letters denote sets with typical elements denoted by lower case letters. Subscripts denote players and unsubscripted letters denote vectors or cartesian products. Thus, $x_i \in X_i$ and $x = (x_1, x_2, ..., x_n) \in X = \times_i X_i$. Also, x_{-i} denotes the vector obtained after the ith component of x has been removed and (x_i', x_{-i}) denotes the vector where the ith component of x has been repaced by x_i'. Finally, $\Delta(X)$ is the set of probability distributions over X.

[10] This ensures that knowledge of one's ex post payoff does not carry any infomation beyond that in the signal. For instance, in Stigler's (1964) model a firm's profits depend only on its own actions (prices) and its own signal (sales).

where $\|\mu - \nu\|_{TV}$ denotes the *total variation* distance between the probability measures μ and ν.[11] It is intuitively clear that when the quality of monitoring is poor, it is hard for players other than i to detect a deviation by i.

Coarse correlated equilibrium The distribution $\alpha \in \Delta(A)$ is a *coarse correlated equilibrium* (CCE) of G if for all i and all $a_i \in A_i$,

$$u_i(\alpha) \geq u_i(a_i, \alpha_{-i})$$

where $\alpha_{-i} \in \Delta(A_{-i})$ denotes the marginal distribution of α over A_{-i} (see Moulin and Vial, 1978).

The notion of a CCE is best understood by contrasting it with the notion of a classical correlated equilibrium (CE). A correlated equilibrium can be thought of as a "mediated solution"—a mediator draws a joint action $a \in A$ from a distribution $\alpha \in \Delta(A)$ and sends a private recommendation to each player. The distribution α constitutes a CE if for every a in the support of α, no player can gain by choosing a different action than the one recommended. A coarse correlated equilibrium α can also be thought of as a "mediated solution but with commitment"—the players have to decide whether or not to "sign on" to the mediated solution without knowing anything other than the distribution α. A CCE involves greater commitment on the part of the players than does a CE—they agree to play according to a joint agreement α knowing nothing else.[12]

The distribution $\alpha \in \Delta(A)$ is an *ε-coarse correlated equilibrium* (ε-CCE) of G if for all i and all $a_i \in A_i$,

$$u_i(\alpha) \geq u_i(a_i, \alpha_{-i}) - \varepsilon \tag{3}$$

Define ε-$CCE(G) = \{u(\alpha) \in \mathcal{F} : \alpha \text{ is an } \varepsilon\text{-}CCE \text{ of } G\}$ to be the set of ε-coarse correlated equilibrium payoffs of G.

Repeated game We will study an infinitely repeated version of G, denoted by G_δ, defined as follows. Time is discrete and indexed by $t = 1, 2, ...$ and in each period t, the game G is played. Payoffs in the repeated game G_δ are discounted averages of per-period payoffs using the common discount factor $\delta \in (0, 1)$. Precisely, if the sequence of actions taken is $(a^1, a^2, ...)$, player i's ex ante expected payoff is $(1 - \delta) \sum_t \delta^t u_i(a^t)$. A (behavioral) *strategy* for player i in the game G_δ is a sequence of functions $\sigma_i = (\sigma_i^1, \sigma_i^2, ...)$ where $\sigma_i^t : A_i^{t-1} \times Y_i^{t-1} \to \Delta(A_i)$. Hence, a strategy determines a player's current, possibly mixed, action as a function of his private history—his own past actions and past signals.

[11] The total variation distance between two probability measures μ and ν on X is defined as $\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in X} |\mu(x) - \nu(x)|$. This is, of course, equivalent to the metric derived from the L^1 norm.

[12] Further discussion along these lines can be found in Moulin et al. (2014).

Repeated game with communication We will also study a version of G_δ, denoted by G_δ^{com}, in which players can communicate with each other after every period by sending *public* messages m_i from a finite set M_i. The communication phase in period t takes place after the signals in period t have been observed. Thus, a strategy of player i in the game G_δ^{com} consists of two sequences of functions $\sigma_i = (\sigma_i^1, \sigma_i^2, ...)$ and $\rho_i = (\rho_i^1, \rho_i^2, ...)$ where $\sigma_i^t : A_i^{t-1} \times Y_i^{t-1} \times M^{t-1} \to \Delta(A_i)$ determines a player's current action as a function of his own past actions, past signals and past messages from *all* the players. The function $\rho_i^t : A_i^t \times Y_i^t \times M^{t-1} \to \Delta(M_i)$ determines a player's current message as a function of his own past and *current* actions and signals as well as past messages from *all* the players. The messages m_i themselves have no direct payoff consequences.

Equilibrium notion We will consider *sequential* equilibria of the two games. For G_δ, the repeated game without communication, the full support condition (1) ensures that the set of sequential equilibrium payoffs coincides with the set of Nash equilibrium payoffs (see Sekiguchi, 1997). In both situations, we suppose that players have access to public randomization devices.

3 Some examples

Before beginning a formal analysis of equilibrium payoffs in the repeated game G_δ and its counterpart with communication, G_δ^{com}, it will be instructive to consider a few examples. The first example illustrates, in the simplest terms, the main result of the paper. The other examples then point to some complexities. A word of warning is in order. All of the following examples have the property that the marginal distributions of players' signals are the same regardless of players' actions. This means that the expected payoff functions $u_i(a)$ cannot be derived from underlying ex post payoff functions $w_i(a_i, y_i)$ and, of course, expected payoffs are *not* observed. The examples have this property only to illustrate some features of the model in the simplest way possible. This is not essential—the examples can easily be amended so that underlying ex post payoff functions, which are observed, exist.

Example 1: Communication is necessary for cooperation. Consider the following prisoners' dilemma as the stage game with expected payoffs:

	c	d
c	$2, 2$	$-1, 3$
d	$3, -1$	$0, 0$

Each player has two possible signals y' and y'' and suppose that the monitoring structure q is

$$q(\cdot \mid cc) = \begin{array}{c|c|c|} & y' & y'' \\ \hline y' & \frac{1}{2}-\varepsilon & \varepsilon \\ \hline y'' & \varepsilon & \frac{1}{2}-\varepsilon \\ \hline \end{array} \qquad q(\cdot \mid \neg\, cc) = \begin{array}{c|c|c|} & y' & y'' \\ \hline y' & \frac{1}{4} & \frac{1}{4} \\ \hline y'' & \frac{1}{4} & \frac{1}{4} \\ \hline \end{array} \tag{4}$$

where $\neg\, cc$ denotes any action profile other than cc.

We argue below that without communication, it is impossible for players to "cooperate"—that is, to play cc—and that the unique equilibrium payoff is $(0, 0)$. With communication, however, it is possible for the players to cooperate (with high probability) and, in fact, attain average payoffs close to $(2, 2)$.

The monitoring structure here has two key features. First, the marginal distributions $q_i(\cdot \mid a)$ are *identical* no matter what action a is played, so that the quality of monitoring η (defined in (2)) is zero. Second, if cc is played, each player's signal is very *informative* about the other player's signal. If something other than cc is played, a player's signal is completely uninformative about the other player's signal.[13] At the end of this example, we derive this monitoring structure from more basic considerations.

Claim 1 *Without communication, cooperation is not possible—for all δ, the unique equilibrium payoff of G_δ is $(0, 0)$.*

Fix any strategy of player 2. Since the marginal distribution on player 2's signals is not affected by what player 1 does, his ex ante belief on what player 2 will play in any future period is also independent of what he plays today. Thus, in any period, player 1 is better off playing d rather than c. Cooperation is impossible.[14]

Claim 2 *With communication, cooperation is possible—given any $\delta > \frac{1}{2}$, there exists an $\bar{\varepsilon}$ such that for all $\varepsilon < \bar{\varepsilon}$, there exists an equilibrium of G_δ^{com} whose payoffs are ε-close to $(2, 2)$.*

Now suppose that players report their signals during the communication phase—that is, $M_i = Y_i$. Consider the following variant of a "trigger strategy": play c in period 1 and in the communication phase, report the signal that was received. In any period t, play c if in all past periods, the reported signals have agreed—that is, if both players reported y' or both reported y''. If the reports disagreed in any past period, play d. In the communication phase, report your signal.

[13] The term "informative" is used in the sense of Blackwell (1951). From player i's perspective, the signals of the other players y_{-i} constitute the "state of nature" and his own signal carries information about this.

[14] The equality of the marginals violates one of Sugaya's (2015) conditions and so his folk theorem does not apply.

To see that these strategies constitute an equilibrium, note first that if all past reports have agreed, and a player has played c in the current period, then there is no incentive to misreport one's signal. Misreporting only increases the probability of triggering a punishment from 2ε to $\frac{1}{2}$ and so there is no gain from deviating during the communication phase.

Finally, if all past reports agreed, a player cannot gain by deviating by playing d. Such a deviation will trigger a punishment with probability $\frac{1}{2}$, no matter what he reports in the communication phase. It is routine to verify that when ε is small, this is not profitable. Each player's payoff in this equilibrium is

$$v = \frac{1-\delta}{1-\delta+2\delta\varepsilon} \times 2$$

which converges to 2 as ε converges to zero.

The monitoring structure in this example can be derived from more basic considerations, as outlined in the introduction. Suppose that there are two equally likely states of nature ω' and ω'' and players get a noisy private signal y' or y'' about the state. Conditional on the state, players' signals are independent. If a player plays c in state ω', then he receives signal y' with probability $1-\lambda > \frac{1}{2}$ and signal y'' with probability λ. Likewise, if a player plays c in state ω'', then he receives y'' with probability $1-\lambda$ and y' with probability λ. On the other hand, if a player plays d, then the two signals are equally likely regardless of state. Thus, cooperating provides information about the state whereas defecting does not. A routine computation shows that the monitoring structure in (4) results for $\varepsilon = \lambda(1-\lambda)$.

Example 2: Cooperation is impossible even with communication The first example exhibited some circumstances in which cooperation was not possible without communication but with communication, it was. Does communication *always* facilitate cooperation? As the next example shows, this is not always the case—the signal structure q matters.

Consider the prisoners' dilemma of Example 1 again but with the following "flipped" signal structure:

$$q(\cdot \mid \neg\, dd) = \begin{array}{c|c|c|} & y' & y'' \\ \hline y' & \frac{1}{4} & \frac{1}{4} \\ \hline y'' & \frac{1}{4} & \frac{1}{4} \\ \hline \end{array} \qquad q(\cdot \mid dd) = \begin{array}{c|c|c|} & y' & y'' \\ \hline y' & \frac{1}{2}-\varepsilon & \varepsilon \\ \hline y'' & \varepsilon & \frac{1}{2}-\varepsilon \\ \hline \end{array}$$

where again, $\neg\, dd$ denotes any action profile other than dd.

The marginal distribution of signals $q_i(\cdot \mid a)$ is, as before, unaffected by players' actions—η is zero again. But now the signal distribution when dd is played is *more* informative than when any other action is played—in fact, the former is completely informative.

Claim 3 *With or without communication, cooperation is not possible—for all δ, the unique equilibrium payoff in both G_δ and G_δ^{com} is $(0,0)$.*

Suppose that there is an equilibrium with communication in which after some history, player 1 is supposed to play c with probability one and report his signal truthfully. Suppose player 1 plays d instead of c and at the communication stage, regardless of his private signal, reports with probability one-half that his signal was y' and with probability one-half that his signal was y''. Now regardless of whether player 2 plays c or d in that period, the joint distribution over player 1's *reports* and player 2's signals is the same as if player 1 had played c—that is, $q(\cdot \mid \neg\, dd)$. Thus, player 1 can deviate and "lie" in a way that his deviation cannot be statistically detected.[15]

Thus, with the "flipped" monitoring structure no cooperation is possible even with communication. A fortiori, no cooperation is possible without communication either. In this example, therefore, communication is unable to facilitate cooperation.

Example 3: Communication is not necessary for cooperation Our final example illustrates the possibility that "full" cooperation is possible without communication even though monitoring is very poor—in fact, non-existent.[16] Consider the following version of "rock-paper-scissors":

	r	p	s
r	10, 10	0, 11	11, 0
p	11, 0	10, 10	0, 11
s	0, 11	11, 0	10, 10

The stage game has a *unique* correlated equilibrium, and hence a unique Nash equilibrium as well, in which players randomize equally among the three actions and results in a payoff of $(7,7)$. Suppose the monitoring structure is:

$$q(\cdot \mid a_1 = a_2) = \begin{array}{c|c|c|c|} & y^r & y^p & y^s \\ \hline y^r & \frac{1}{3} & 0 & 0 \\ \hline y^p & 0 & \frac{1}{3} & 0 \\ \hline y^s & 0 & 0 & \frac{1}{3} \\ \hline \end{array} \qquad q(\cdot \mid a_1 \neq a_2) = \begin{array}{c|c|c|c|} & y^r & y^p & y^s \\ \hline y^r & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \hline y^p & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \hline y^s & \frac{1}{9} & \frac{1}{9} & \frac{1}{9} \\ \hline \end{array}$$

so that if players coordinate on the same action, then the signals are perfectly informative; otherwise, they are uninformative. Once again $\eta = 0$ since the marginal distributions of signals are not affected by players' actions.

Claim 4 *Cooperation is possible without communication—for any $\delta \geq \frac{3}{11}$, there exists an equilibrium of G_δ with a payoff of $(10,10)$.*

In what follows, we will say that the players are "coordinated" if they take the same action and so the resulting signal distribution is $q(\cdot \mid a_1 = a_2)$. Otherwise, they are said to be "miscoordinated."

[15] This argument can be extended to include randomized strategies. Details can be obtained from the authors.

[16] This trivally holds in games where there is an efficient one-shot Nash equilibrium, of course. In this example that is not the case.

Consider the following strategy: in period 1, play r. In period t, play action $a \in \{r, p, s\}$ if the signal received in the last period was y^a.

The average payoff from this strategy is clearly 10 since the players are always coordinated. Now suppose player 1 deviates once from the prescribed strategy and then reverts back to it.

Player 1's immediate payoff from the deviation is 11. But the deviation also causes the players to become miscoordinated. So no matter what signal player 1 receives, player 2 is equally likely to play each of his actions. As a result, once the players are miscoordinated, the continuation payoff is

$$w = \tfrac{1}{3}10 + \tfrac{1}{3}\left((1-\delta)\,11 + \delta w\right) + \tfrac{1}{3}\left((1-\delta)\,0 + \delta w\right)$$

This is because with probability $\frac{1}{3}$, the players will become coordinated again in the next period and then remain coordinated thereafter. With probability $\frac{1}{3}$, they will remain miscoordinated, player 1 will get 11 and then the continuation payoff w; with probability $\frac{1}{3}$, he will get 0 and then w again. Thus the continuation payoff after miscoordination is

$$w = \frac{21 - 11\delta}{3 - 2\delta}$$

The original deviation is not profitable as long as

$$(1-\delta)\,11 + \delta w \leq 10$$

and this holds as long as $\delta \geq \frac{3}{11}$. The one-deviation principle for games with private monitoring (Mailath and Samuelson, 2006, p. 397), then ensures that the prescribed strategies constitute an equilibrium. The example thus demonstrates that in a repeated setting, players can sometimes achieve outcomes far superior to the set of correlated equilibria of the stage game even though there is zero monitoring.

4 Equilibrium without communication

In this section, we develop a method to bound the set of equilibrium payoffs in games with private monitoring. We will show that the set of equilibrium payoffs of the *repeated* game without communication G_δ is contained within the set of ε-coarse correlated equilibrium payoffs of the *stage* game G. In our result, we give an explicit formula for ε involving (a) the discount factor; and (b) the quality of monitoring (as defined in (2)).

The main result of this section is:

Proposition 4.1

$$NE\,(G_\delta) \subseteq \varepsilon\text{-}CCE\,(G)$$

where $\varepsilon = 2\frac{\delta^2}{1-\delta}\eta \times \|u\|_\infty$.

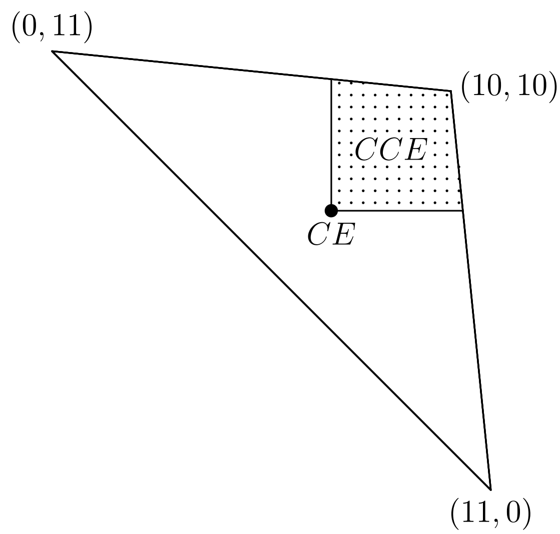

Figure 1: Rock-Scissors-Paper: $CE = (7,7)$ and $CCE =$ shaded area

The proposition is of independent interest because repeated games with private monitoring do not have a natural recursive structure and so a characterization of the set of equilibrium payoffs seems intractable. So one is left with the task of finding effective bounds for this set. Proposition 4.1 provides such a bound and one that is easy to compute explicitly: the set ε-CCE is defined by $\sum_i |A_i|$ linear inequalities. Moreover, the bound does not use any detailed information about the monitoring structure—it depends only on the monitoring quality parameter η.[17]

Proposition 4.1 relates Nash equilibria of the repeated game to coarse correlated equilibria of the one-shot game. It is clear that the signals that players receive result in their actions being correlated and so the relationship to some "mediated solution" is not unnatural. To see why CCE is the right notion, consider a monitoring technology for which $\eta = 0$, that is, a situation in which player i's actions do not affect the distribution of signals of other players. This means, of course, that any deviation by player i will go undetected. Consider a repeated game equilibrium strategy which is *stationary* and so results in a fixed distribution of actions $\alpha \in \Delta(A)$ in every period. Because deviations cannot be detected, each player i can guarantee that his or her payoff is $\max_{a_i} u_i(a_i, \alpha_{-i})$ which implies immediately that α must be a coarse correlated equilibrium. The formal proof below allows for both $\eta > 0$ as well as non-stationary strategies.

The "ε-coarse correlated equilibrium" in the statement cannot simply be replaced with "ε-correlated equilibrium." Precisely, if the set of ε-correlated equilibrium payoffs of G is denoted by ε-$CE(G)$, then the statement $NE(G_\delta) \subseteq \varepsilon$-$CE(G)$ is false for the

[17] Sugaya and Wolitzky (2017) show that the set of equilibrium payoffs with perfect monitoring and a meditor is a bound for large enough δ. Unlike ours, their bound is independent of the quality of monitoring. Our method results in tighter upper bounds to payoffs when the quality of monitoring is poor.

same value of ε as above. For instance, in Example 3 the set of correlated equilibrium payoffs is a singleton $CE(G) = \{(7,7)\}$ while the set of coarse correlated equilibria is as depicted in Figure 1. For the monitoring structure in Example 3, $\eta = 0$ and hence $\varepsilon = 0$ as well. But for $\delta \geq \frac{3}{11}$, repeated game has an equilibrium payoff of $(10, 10)$ which is in $CCE(G)$ but not in $CE(G)$.

Sketch of Proof We indicated how the CCE bound arises naturally when there is zero monitoring ($\eta = 0$) and the equilibrium strategy of the repeated game is stationary. We now show how the argument is extended to permit both $\eta > 0$ and non-stationary strategies. So consider strategy profile σ of the repeated game with a payoff $v(\sigma)$ that is not an ε-CCE payoff in the one-shot game. Suppose that some player i deviates to a strategy $\overline{\sigma}_i$ in which i chooses $\overline{a}_i$ in every period regardless of history—that is, $\overline{\sigma}_i$ consists of a *permanent* deviation to $\overline{a}_i$. We decompose the (possible) gain from such a deviation into two bits. Consider a fictitious situation in which the players $j \neq i$ are replaced by a non-responsive *machine* that, in every period, and regardless of history, plays the ex ante distribution $\alpha^t_{-i} \in \Delta(A_{-i})$ that would have resulted from the candidate strategy σ. In the fictitious situation, player i's deviation is unpunished in the sense that the machine continues to play as if no deviation had occurred. Note that this lack of response is what would have occurred if η were zero. We can then write

$$\underbrace{v_i(\overline{\sigma}_i, \sigma_{-i}) - v_i(\sigma)}_{\text{Gain from deviation}} = \underbrace{v_i(\overline{\sigma}_i, \sigma_{-i}) - v_i(\overline{\sigma}_i, \alpha_{-i})}_{\text{Loss from punishment}} + \underbrace{v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\sigma)}_{\text{Gain when unpunished}} \tag{5}$$

The first component on the right-hand side represents the payoff difference from facing the real players $j \neq i$ versus facing the non-responsive machine. If σ_{-i} is an effective deterrent to the permanent deviation then this should be negative and in Lemma 4.1 we calculate a lower bound to this loss. The second component is the gain to player i when his permanent deviation goes unpunished. As we will show below in Lemma 4.2, this gain can be related to the coarse correlated equilibria of the one-shot game (see (3)). The lemma establishes that a permanent deviation, which is, of course, stationary, is best deterred by a stationary distribution.

Non-responsive strategies We begin with a formal definition of the *non-responsive* strategy played by the fictitious "machine." Given a strategy profile σ, the induced ex ante distribution over A in period t is

$$\alpha^t(\sigma) = E_\sigma\left[\Pi_j \sigma^t_j\left(h^{t-1}_j\right)\right] \in \Delta(A)$$

and the corresponding marginal distribution over A_{-i} in period t is

$$\alpha^t_{-i}(\sigma) = E_\sigma\left[\Pi_{j\neq i} \sigma^t_j\left(h^{t-1}_j\right)\right] \in \Delta(A_{-i})$$

where the expectation is defined by the probability distribution over $t-1$ histories determined by σ. Note that α_{-i} depends on the whole strategy *profile* σ and not just on the strategies σ_{-i} of players other than i. Note also that because players' histories are correlated, it is typically the case that $\alpha_{-i}^t(\sigma) \notin \Pi_{j\neq i}\Delta(A_j)$. Given σ, let $\alpha_{-i}(\sigma)$ denote the (correlated) strategy of players $j \neq i$ in which they play $\alpha_{-i}^t(\sigma)$ in period t following *any* $t-1$ period history. The strategy α_{-i}, which is merely a sequence $\{\alpha_{-i}^t\}$ of joint distributions in $\Delta(A_{-i})$, replicates the *ex ante* distribution of actions of players $j \neq i$ resulting from σ but is non-responsive to histories.

We now proceed to decompose the gain from a permanent deviation.

4.1 Loss from punishment

In this subsection we provide a bound on the absolute value of $v_i(\overline{\sigma}_i, \sigma_{-i}) - v_i(\overline{\sigma}_i, \alpha_{-i})$, the difference in payoffs between being punished by strategy σ_{-i} of the real players $j \neq i$ versus not being punished by the fictitious machine. It is clear that the magnitude of this difference depends crucially on how responsive σ_{-i} is compared to the α_{-i} and this in turn depends on how well the players $j \neq i$ can detect i's permanent deviation.

We show below that this loss can in fact be bounded by a quantity that is a positive linear function of η. Moreover, the bound is increasing in δ. The following result provides an exact formula for the trade-off between the quality of monitoring and the discount factor.

Lemma 4.1 *Suppose i plays $\overline{a}_i$ always. The difference in i's payoff when others play σ_{-i} versus when they play the non-responsive strategy α_{-i} derived from σ satisfies*

$$|v_i(\overline{\sigma}_i, \sigma_{-i}) - v_i(\overline{\sigma}_i, \alpha_{-i})| \leq 2\frac{\delta^2}{1-\delta}\eta \times \|u\|_\infty$$

Proof. See Appendix A. ■

4.2 Gain when unpunished

We now relate the second component in (5) to the ε-coarse correlated equilibria of the one-shot game (see (3)). We begin by characterizing the set of ε-CCE payoffs.

For any $v \in \mathcal{F}$, define

$$\Theta(v) \equiv \min_{\beta\in\Delta(A)} \max_i \max_{\overline{a}_i} \left[u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)\right]$$

subject to

$$u(\beta) = v$$

where $\beta_{-i} \in \Delta(A_{-i})$ denotes the marginal distribution of β over A_{-i}.

In words, no matter how the payoff v is achieved via a correlated action, at least one player can gain at least $\Theta(v)$ by deviating. It is easy to see that $v \in \varepsilon$-$CCE(G)$

if and only if $\Theta(v) \leq \varepsilon$. This is because $\Theta(v) \leq \varepsilon$ is the same as: there exists a $\beta \in \Delta(A)$ satisfying $u(\beta) = v$ such that

$$\max_i \max_{\overline{a}_i \in A_i} \left[u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)\right] \leq \varepsilon$$

and this is equivalent to $v \in \varepsilon\text{-}CCE(G)$.

Note that $\Theta(v)$ is also the value of an artificial two-person zero-sum game Γ in which player I ("deviator") chooses a pair $(i, \overline{a}_i)$ and player II ("mediator") chooses a joint distribution $\beta \in \Delta(A)$ such that $u(\beta) = v$. The idea is that the mediator chooses a joint distribution and the deviator chooses a player and a pure strategy for that player as a deviation. The payoff to player I is then $u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)$. The fact that $\Theta(v) \leq 0$ is the same as $v \in CCE(G)$ is analogous to a result of Hart and Schmeidler (1989) on correlated equilibria.

The following important result shows that the function Θ, which measures the *static* incentives to deviate, also measures the *dynamic* incentives to *permanently* deviate from a non-responsive strategy. Formally, consider a dynamic analogue Γ_δ of the two-person zero-sum game Γ outlined above. In Γ_δ, player I ("deviator") chooses a pair $(i, \overline{\sigma}_i)$ where $\overline{\sigma}_i$ denotes the constant sequence $\overline{a}_i$ (a permanent deviation) and player II ("mediator") chooses a non-responsive strategy—that is, a *sequence* of joint distributions $\alpha \in \Delta(A)^\infty$—such that its discounted average $v(\alpha) = v$. The payoff to player I in Γ_δ is $v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\alpha)$. The lemma shows that the value of Γ, that is, $\Theta(v)$, is also the value of Γ_δ. This relies on the fact that the deviation is permanent (stationary). Since the maximized payoff function of player I is convex in player II's strategy, player II optimal response is stationary as well. This last step resembles the familiar "consumption smoothing" argument.

Lemma 4.2

$$\Theta(v) = \min_{\alpha \in \Delta(A)^\infty} \max_i \max_{\overline{\sigma}_i} \left[v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\alpha)\right]$$

subject to the constraint that the discounted average payoff from α,

$$v(\alpha) = v$$

where $v_i(\overline{\sigma}_i, \alpha_{-i})$ *is* i*'s payoff when he plays* $\overline{a}_i$ *always and others play the non-responsive strategy* $\alpha_{-i} = \left(\alpha^1_{-i}, \alpha^2_{-i}, ...\right) \in \Delta(A_{-i})^\infty$ *derived from* $\alpha = (\alpha^1, \alpha^2, ...) \in \Delta(A)^\infty$.

Proof. It is clear that $\Theta(v)$ is at least as large as the right-hand side of the equality above. This is because the set of strategies available to player II in Γ_δ includes all stationary strategies and the latter are equivalent to all strategies in Γ. The set of strategies available to player I in Γ and Γ_δ are the same.

Given a sequence $\alpha = (\alpha^1, \alpha^2, ...) \in \Delta(A)^\infty$, let $v^t = u(\alpha^t)$ be the ex ante payoffs in period t. Then, if $v(\alpha) = v$ we have $(1-\delta)\sum_{t=1}^\infty \delta^t v^t = v$.

For any payoff vector $w \in \mathcal{F}$, define $\theta_i(w, \overline{a}_i) = \min_{\beta \in \Delta(A)} \left[u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)\right]$ subject to $u(\beta) = w$. Then, for for all i and all $\overline{a}_i$,

$$\begin{aligned} v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\alpha) &= (1-\delta)\sum_{t=1}^{\infty}\delta^t\left[u_i\left(\overline{a}_i, \alpha_{-i}^t\right) - u_i\left(\alpha^t\right)\right] \\ &\geq (1-\delta)\sum_{t=1}^{\infty}\delta^t\theta_i\left(v^t, \overline{a}_i\right) \end{aligned} \tag{6}$$

where the second inequality follows from the definition of θ_i.

Now since $\theta_i(\cdot, \overline{a}_i)$ is convex[18], it is the case that a solution to the problem:

$$\min_{v^t}(1-\delta)\sum_{t=1}^{\infty}\delta^t\theta_i\left(v^t, \overline{a}_i\right)$$

subject to

$$(1-\delta)\sum_{t=1}^{\infty}\delta^t v^t = v$$

is to set $v^t = v$ for all t. Thus, we have that

$$(1-\delta)\sum_{t=1}^{\infty}\delta^t\theta_i\left(v^t, \overline{a}_i\right) \geq \theta_i(v, \overline{a}_i)$$

which when combined with (6) yields that for all i and $\overline{a}_i$,

$$v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\alpha) \geq \theta_i(v, \overline{a}_i) = \min_{\beta: u(\beta) = v}\left[u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)\right]$$

This implies that

$$\min_{\alpha: v(\alpha) = v}\left[v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\alpha)\right] \geq \min_{\beta: u(\beta) = v}\left[u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)\right]$$

and thus

$$\max_{\pi \in \Delta(\cup A_j)}\min_{\alpha: v(\alpha) = v} E_\pi\left[v_i(a_i, \alpha_{-i}) - v_i(\alpha)\right] \geq \max_{\pi \in \Delta(\cup A_j)}\min_{\beta: u(\beta) = v} E_\pi\left[u_i\left(a_i, \beta_{-i}\right) - u_i(\beta)\right]$$

Applying the minmax theorem (Sion, 1958) in the game Γ_δ on the left-hand side and in the game Γ on the right-hand side, we obtain

$$\begin{aligned} \min_{\alpha: v(\alpha) = v}\max_i\max_{\overline{a}_i} v_i(\overline{\sigma}_i, \alpha_{-i}) - v_i(\alpha) &\geq \min_{\beta: u(\beta) = v}\max_i\max_{\overline{a}_i}\left[u_i\left(\overline{a}_i, \beta_{-i}\right) - u_i(\beta)\right] \\ &= \Theta(v) \end{aligned}$$

■

[18] The convexity of $\theta_i(\cdot, \overline{a}_i)$ is a consequence of the fact that u is linear in α.

4.3 Payoff bound

With Lemmas 4.1 and 4.2 in hand, we can now complete the proof of the result (Proposition 4.1) that the set of Nash equilibrium payoffs of the repeated game is contained in the set of ε-coarse correlated equilibrium payoffs of the one-shot game.

Suppose σ is a strategy profile in G_δ such that $v \equiv v(\sigma) \notin \varepsilon\text{-}CCE(G)$ for $\varepsilon = 2\frac{\delta^2}{1-\delta}\eta \times \|u\|_\infty$. Then we know that $\Theta(v) > \varepsilon$. Lemma 4.2 implies that

$$\min_{\alpha \in \Delta(A)^\infty} \max_i \max_{\overline{\sigma}_i} \left[v_i\left(\overline{\sigma}_i, \alpha_{-i}\right) - v_i(\alpha)\right] > \varepsilon$$

and so

$$\max_i \max_{\overline{\sigma}_i} \left[v_i\left(\overline{\sigma}_i, \alpha_{-i}\right) - v_i(\sigma)\right] > \varepsilon$$

where α is the non-responsive strategy derived from σ as above. Thus, there exists a player i and a permanent deviation for that player such that $v_i\left(\overline{\sigma}_i, \alpha_{-i}\right) - v_i(\sigma) > \varepsilon$. Applying Lemmas 4.1 and 4.2 we have

$$\begin{aligned} v_i\left(\overline{\sigma}_i, \sigma_{-i}\right) - v_i(\sigma) &= v_i\left(\overline{\sigma}_i, \sigma_{-i}\right) - v_i\left(\overline{\sigma}_i, \alpha_{-i}\right) + v_i\left(\overline{\sigma}_i, \alpha_{-i}\right) - v_i(\sigma) \\ &> -2\frac{\delta^2}{1-\delta}\eta \|u\|_\infty + \varepsilon \\ &= 0 \end{aligned}$$

Thus, σ is not a Nash equilibrium of G_δ. This completes the proof of Proposition 4.1.

4.4 Bound Comparison

The bound in Proposition 4.1 is related to but (weakly) tighter than the bound obtained in our earlier work (Awaya and Krishna, 2016), which is called the Ψ-bound as it depends on a function Ψ defined there. The Ψ-bound was obtained by considering a permanent deviation to a *single* deviating action $\overline{a}_i = a_i^{BR}$, a one-period best response to a^* which uniquely achieves an efficient payoff u^*. The new CCE-bound, derived in this section, also considers a permanent deviation but both the identity of the deviator i and the deviating action $\overline{a}_i$ are chosen *optimally.*

In the prisoners' dilemma, there is no difference in the bounds but in many other games, the resulting improvement can be substantial. For instance, in the game,

	b	c	d
b	$8, 8$	$0, 0$	$-10, 7$
c	$0, 0$	$8, 8$	$-10, 10$
d	$10, -10$	$7, -10$	$0, 0$

the Ψ-bound is trivial. It says merely that all equilibrium payoffs of the repeated game satisfy $\frac{1}{2}v_1 + \frac{1}{2}v_2 \leq 8$ (the symmetric efficient payoff in the one-shot game). The reason the earlier Ψ-bound is trivial is that there is an efficient mixed action which places probability $\frac{1}{2}$ on $(b, b,)$ and probability $\frac{1}{2}$ on (c, c) resulting in a payoff of $(8, 8)$. But when (b, b) is played, player 1 has the incentive to deviate to d and when (c, c) is played, player 2 has the incentive to deviate to d. The Ψ-bound does not allow the deviator be chosen depending on which of (b, b) and (c, c) is played—the single deviator must be chosen once and for all. Thus, the Ψ-bound under-estimates the incentive to deviate. In this game, however, the unique CCE payoff is $(0, 0)$ and the highest symmetric ε-CCE payoff is $(16\varepsilon, 16\varepsilon)$. Proposition 4.1 now implies that equilibrium payoffs of the repeated game satisfy $\frac{1}{2}v_1 + \frac{1}{2}v_2 \leq 16\varepsilon$.

The improvement in the bound can be calculated in any game (both the function Ψ and the set of CCE payoffs can be computed via linear programs). To get some sense of the improvement in a game of economic interest, consider a duopoly in which firms compete by setting prices for differentiated products. This was the model studied in Awaya and Krishna (2016). Suppose that the firms' demands are linear in prices: for $i = 1, 2$ and $j \neq i$,

$$d_i = A - bp_i + p_j$$

where $b > 1$. Note that since $b > 1$, the "own-price effect" on a firm's demand is greater than the "cross-price effect." In this example, the difference in total profits given by the two bounds when there is zero monitoring ($\eta = 0$) can be obtained in closed form:

$$\Psi\text{-bound} \; - \; CCE\text{-bound} = A^2 \left(\frac{1}{8} \frac{1}{b^2 (2b-1)^2} \frac{4b-1}{b-1} \right) \tag{7}$$

The Ψ-bound for this game can be calculated using the derivation in Awaya and Krishna (2016, p. 305). In the case of linear demand, the set of CCE of this game coincides with the unique Nash equilibrium (see Gérard-Varet and Moulin, 1978) and is easily calculated.

To get a sense of the improvement, suppose $b = 2$. Then the Ψ-bound says that when the monitoring is poor in any equilibrium of the repeated game, the total profits of the two firms cannot exceed approximately 44% of the gap between total monopoly profits and total Nash profits. But the CCE-bound is much tighter. It says that with poor monitoring, total profits in any equilibrium cannot exceed approximately 1% of the gap.

5 Equilibrium with communication

In what follows, we will consider efficient actions a^* that Pareto dominate some Nash equilibrium of the stage game, that is, $u(a^*) \gg u(\alpha^N)$ where $\alpha^N \in \times_i \Delta(A_i)$ is a (possibly mixed) Nash equilibrium of the stage game.[19] In Section 5.1 below, we will display a particular strategy profile for the game with communication. Then in

[19] This condition can be easily weakened to require only that $u(a^*)$ Pareto dominate some convex combination of one-shot Nash equilibrium payoffs.

Section 5.2, we will identify conditions on the signal structure q and the discount factor δ that guarantee that the profile constitutes an equilibrium which is "nearly" efficient (see Proposition 5.1 below). We emphasize that our result is not a "folk theorem." In the latter, the signal structure is held fixed and the discount factor is raised sufficiently so that any feasible outcome can arise in equilibrium. In our result, the discount factor is held fixed (perhaps at some high level) and the monitoring structure is varied so that efficient outcomes can be sustained in equilibrium.

In what follows, it will convenient to assume that all players have the same set of signals—that is, for all i and j, $Y_i = Y_j$ and without loss of generality, we will suppose that for all i, the set of signals $Y_i = \{1, 2, ..., K\}$.

5.1 A simple equilibrium

We will now construct a nearly efficient equilibrium with communication. Define Y^D to be the set of (diagonal) signal profiles such that $y_1 = y_2 = ... = y_n$, that is, profiles in which all players' signals are the same.

The proposed equilibrium strategy resembles a trigger strategy and is very simple: In period 1, play a_i^* and report the signal received truthfully. In any period $t > 1$, play a_i^* and report the signal y_i^t truthfully if all players have reported the *same* signal in *all* past periods, that is, if all players have reported the same signal in every period $s < t$, that is, if for all $s < t$, $y^s \in Y^D$. Otherwise, play the one-shot Nash action α_i^N and report y_i^t truthfully.

The strategy thus requires that all players *unanimously* agree on a signal in order to continue cooperating. Any dissent results in infinite punishment. Another way to write this, useful for later comparisons, is that the strategy calls on players to cooperate in period t if and only if $\mathbb{I}_{Y^D}(y^s) = 1$ in all periods $s < t$, where $\mathbb{I}_{Y^D}$ is the *indicator function* of the set of diagonal profiles Y^D.

Fix the set of signal distributions $\{q(\cdot \mid a) : a \neq a^*\}$ and suppose that for all $a \neq a^*$, $q^*(\cdot \mid a)$ has full support.

As a first step, consider a signal distribution $q^* = q(\cdot \mid a^*)$ that is degenerate on signal profiles in which all the players' signals agree and assigns positive probability to all such profiles. Formally, $q^*(y) > 0$ if and only if $y_1 = y_2 = ... = y_n$. We will call such a distribution *perfectly informative* since any player's signal provides perfect information about others' signals. Thus, the terminology is consistent with that of Blackwell (1951).

First, suppose that all players follow the suggested strategies. Then the payoff of player i is $u_i(a^*)$ since punishments are never triggered. We will now argue that if no one has deviated until now, then no player has any incentive (i) to misreport his signal; or (ii) to deviate to another action.

Suppose that the suggested strategies are being played and we are in a situation in which in all past periods, players' reports have agreed. If player i plays a_i^* in period t and then receives the signal y_i she is sure that all other players' signals are the same as hers. Reporting y_i truthfully is then strictly better than reporting any other signal since doing the latter is sure to trigger a punishment. Thus, no player has the incentive to lie along the equilibrium path. After a disagreement, the play is independent of the reports.

Finally, provided that the discount factor is high enough, no player has any incentive to deviate if we are along the equilibrium path. This is because following any deviation to $a_i \neq a_i^*$, the distribution $q\left(\cdot \mid a_i, a_{-i}^*\right)$ will assign positive probability to all signal profiles and in particular to profiles where not all signals agree. Thus any deviation will trigger punishment with positive probability and even by optimally tailoring her report following a deviation, a player cannot reduce this to zero. Provided the discount factor is high enough, no deviation will be profitable and all the incentives can be made strict.

Thus we have argued that if q^* is perfectly informative and there is a high enough discount factor so that the repeated game with communication has an equilibrium which is fully efficient. Now suppose that q^* is nearly informative in the sense that it is close to a perfectly informative distribution q^0. Then by continuity we obtain

Proposition 5.1 *Fix $\{q(\cdot \mid a) : a \neq a^*\}$. There exists a $\underline{\delta}$ such that for all $\delta > \underline{\delta}$ and for all perfectly informative $q^0 \in \Delta(Y)$, there exists a $\overline{\gamma}$ such that if $\|q^0 - q^*\|_{TV} < \overline{\gamma}$, there is a nearly efficient equilibrium of the game with communication.*

Note the distributions $q(\cdot \mid a)$ for $a \neq a^*$ are fixed and it is only $q(\cdot \mid a^*)$ that is required to be nearly informative.[20]

Some shortcomings of the equilibrium The equilibrium constructed above, however, has some shortcomings. First, with the suggested strategies, the likelihood that cooperation may break down even though no one has deviated may be substantial—players will revert to noncooperation whenever there is *any* discrepancy in the reported signals. Such discrepancies will occur, with probability γ, even if all players have conformed to the equilibrium strategy. Second, the strategies guarantee truthtelling only when $\|q^0 - q^*\|_{TV}$ is small, that is, when q^* is very close to being perfectly informative.

The first deficiency is particularly acute when the number of signals or the number of players is large.

1. To see this in the case of a large number of signals, suppose that there are two players and suppose, as a limiting case, that there is a continuum of signals in $[0, 1]$. The probability that the signals of the two players are the same is clearly zero and so the probability of continuing cooperation is zero as well. A discrete approximation to the continuous distribution will have the property

[20] This means that setting we study is not one with "almost-public signals" as in Mailath and Morris (2002) or Hörner and Olszewski (2009).

that the probability of cooperation is small. A similar phenomenon arises with a large number of players (for a fixed signal generating process). As an example, suppose the set of signals for all n players is $Y_i = \{1, 2\}$ and players' signals are independent conditional on an underlying state $\omega \in \{\omega^1, \omega^2\}$. Given $\omega = \omega^k$, each player independently receives the signal k with probability $1 - \lambda > 0.5$ and the other signal $3 - k$ with probability λ. Routine calculations show that the probability of continuing to play a^* if no one has deviated so far is $(1 - \lambda)^n + \lambda^n$, which decreases rapidly to zero as the number of players n increases.

2. The magnitude of the second deficiency can be seen by considering the following example. Suppose that $Y_i = \{1, 2, 3\}$ and the joint distribution of signals is: for $\varepsilon \in (0, \frac{1}{15}]$,

$$q^* = \begin{array}{c|c|c|c|} & 1 & 2 & 3 \\ \hline 1 & \frac{1}{2}(1-15\varepsilon) & 2\varepsilon & 3\varepsilon \\ \hline 2 & 2\varepsilon & \varepsilon & 2\varepsilon \\ \hline 3 & 3\varepsilon & 2\varepsilon & \frac{1}{2}(1-15\varepsilon) \\ \hline \end{array} \tag{8}$$

For this joint distribution, the indicator function (or unanimity rule) does not induce truth-telling no matter how small the value of ε is. This is because for all ε, $q^*(2, 2) < q^*(1, 2)$ and so a player receiving the signal $y_i = 2$ would be better off reporting $z_i = 1$. The reason is that if we fix a q^0 satisfying $q^0(y) > 0$ if and only if all signals in y agree, then for any ε, the total variation distance between q^* and q^0 is at least $\frac{1}{2}q^0(2, 2)$.

We now construct a different equilibrium that is not subject to these deficiencies.

5.2 Another equilibrium

As before, the new strategy profile is also similar to the trigger strategy above but with a different "trigger." The play in any period is governed by a state variable that takes on two values—"normal" and "punishment." The players' strategies depend only on the state: if the state in period t is "normal," play a_i^*; if the state is "punishment," play the one-shot Nash action α^N. The state transitions from period t to $t + 1$ are determined solely by the players' reports of their signals in period t.

The initial state, in period 1, is "normal." If the state is "normal" in period t, and the players report signals y at the end of the period, then the probability that the state remains "normal" in period $t + 1$ is $p^*(y)$ and how the reports y are aggregated to yield $p^*(y)$ is specified in detail below. The "punishment" state is absorbing—if the state is "punishment" in any period, it remains so in every subsequent period. Unlike the indicator function used in the previous subsection, here the probability of punishment is not just 0 and 1. Following a vector of reports, the probability of punishment could be positive but not one.

It is simplest to suppose that following the exchange of signals, for every $y \in Y$, players have access to a public correlation device (a biased coin) which indicates whether the state is "normal" (with probability $p^*(y)$) or "punishment" (with probability $1 - p^*(y)$). In fact, the public correlation device can be replicated by face-to-face communication by making use of jointly controlled lotteries, as in Aumann and Maschler (1995).

Specifically, for all i, consider the following strategy (σ_i^*, ρ_i^*) in the repeated game with communication. The set of messages is the same as the set of signals, that is, $M_i = Y_i$.

The actions chosen according to σ_i^* are as follows:

- In period 1, choose a_i^*.
- In any period $t > 1$, if the state is "normal," choose $a_i^t = a_i^*$; otherwise, choose α_i^N.

The messages sent according to ρ_i^* are as follows:

- In any period $t \geq 1$, if the action chosen $a_i^t = a_i^*$, then report $m_i^t = y_i^t$.
- In any period $t \geq 1$, if the action chosen $a_i^t = a_i \neq a_i^*$ and the signal received is y_i then report $m_i^t \in \arg\max_{z_i \in Y_i} E\left[p^*(z_i, \widetilde{y}_{-i}) \mid y_i\right]$ where the expectation is taken with respect to the distribution $q\left(\cdot \mid (a_i, a_{-i}^*)\right)$.

We will show that there exists a score function $p^*(y)$ with the following key properties:

1. along the equilibrium path, p^* induces truthful reporting of signals;
2. if a player deviates from a_i^* in the normal state, then the expected value of p^* falls (the probability of going to the punishment state increases) regardless of his or her report;
3. if the signals are close to being perfectly informative along the equilibrium path, then the probability of punishment is small.

The first two properties are required for the strategies outlined above to constitute an equilibrium. The third property guarantees that the equilibrium is "nearly" efficient. Figure 2 is a schematic depiction of the play resulting from the given strategies.

Until now we have made no assumptions about the signals and their distribution. In particular, the bound on payoffs without communication, obtained in the previous section, applies without any specific assumptions about the signal structure. In this section, however, we will use some specific features of the signals—that the distribution $q^* = q(\cdot \mid a^*)$ is "positively correlated" in a sense made precise below.

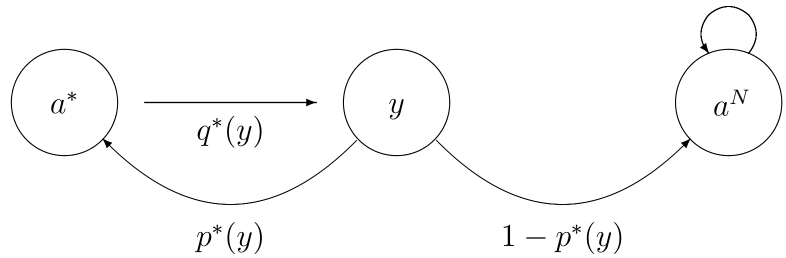

Figure 2: Equilibrium strategy

Recall that the set of signals of any player is $Y_i = \{1, 2, ..., K\}$ and that this has a natural order. Denote by E^* all expectations with respect to q^*. We will suppose that for all i and j :

$$E^* [\widetilde{y}_j \mid y_i] \tag{9}$$

is a *strictly increasing* function of y_i.[21]

We now study circumstances in which the strategies given above constitute an equilibrium. This involves two steps. First, we will construct a particular score function p^* and show that a player who conforms to the equilibrium strategy has no incentive to lie in the communication stage. Next, we will show that no player has any incentive to deviate from the given strategy determined by the particular score function p^*.

5.2.1 Optimality of reports

We now turn to the construction of a score function p^* which will be used to support the equilibrium with communication. This requires that players have the incentive to report their signals truthfully.

Given $q^* \in \Delta(Y)$, consider a payoff function $P^* : Y \to \mathbb{R}$, which is common for all players, defined by

$$P^*(y) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j \neq i} y_i y_j - \sum_{i=1}^{n} \sum_{x_i < y_i} E^* \left[\textstyle\sum_{j \neq i} \widetilde{y}_j \mid x_i \right] - \frac{1}{2} \sum_{i=1}^{n} E^* \left[\textstyle\sum_{j \neq i} \widetilde{y}_j \mid y_i \right] \tag{10}$$

where E^* is the expectation operator with respect to the distribution q^*. We will argue that if all players have the *same* payoff $P^*(y)$ when the reports are y, then they have the incentive to report truthfully. The function P^* is thus a potential function as in Monderer and Shapley (1996). Some understanding of the function may be obtained by considering the case of two players and suppose player 1's true signal is $y_i = k$.

[21] This would be guaranteed, for instance, if the distribution q^* were (strictly) affiliated.

The difference in P^* if player 1, say, reports $y_i = k+1$ rather than $y_i = k$, is then

$$\Delta P_1^* (k) = y_2 - \frac{1}{2} \left(E^* [\widetilde{y}_2 \mid \widetilde{y}_1 = k] + E^* [\widetilde{y}_2 \mid \widetilde{y}_1 = k+1] \right)$$

and this represents the error in player 1's "forecast"—the conditional expectation—of player 2's signal averaged over $\widetilde{y}_1 = k$ and $\widetilde{y}_1 = k+1$. Player 1's expected gain from reporting $k+1$ versus k conditional on his true signal is

$$E [\Delta P_1^* (k) \mid \widetilde{y}_1 = k] = \frac{1}{2} \left(E^* [\widetilde{y}_2 \mid \widetilde{y}_1 = k] - E^* [\widetilde{y}_2 \mid \widetilde{y}_1 = k+1] \right)$$

and since the conditional expectations—the "forecasts"—are increasing in his own signal, this is not profitable if the payoffs are given by P^*. Similarly, reporting $k-1$ is not profitable either and condition (9) guarantees that large deviations are unprofitable as well. Proposition B.1 in Appendix B shows formally that if $q^* \in \operatorname{int} \Delta (Y)$ satisfies (9), then each player has a strict incentive to tell the truth if others are doing so.

The normalized version of P^*, called the *score* function, is

$$p^* (y) = \frac{P^* (y) - \min P^*}{\max P^* - \min P^*} \tag{11}$$

where $\min P^*$ (resp. $\max P^*$) denotes the minimum (resp. maximum) value of P^* over Y. Since Y is finite, the minimum and maximum exist and because of Proposition B.1, $\max P^* > \min P^*$. Thus, for all y, $p^* (y) \in [0, 1]$. Note that p^* is just an affine transformation of P^* and so players' incentives are not affected by this renormalization.

5.2.2 Optimality of actions

Having shown that if a player has not deviated, then it is optimal to report truthfully, we now turn to the optimality of actions. We first show that it is optimal to follow the suggested strategies in the limit case where the distribution $q^* = q(\cdot \mid a^*)$ is perfectly informative—a player's signal gives precise information about others' signals. We then show that the same is true if q^* is near a perfectly informative distribution.

A limit case Recall that a distribution $q \in \Delta (Y)$ is *perfectly informative* if $q(y) > 0$ if and only if $y_1 = y_2 = ... = y_n$. When the distribution is perfectly informative, then every player's individual signal provides perfect information about the signals of the other players.

Now consider the function,

$$P^0 (y) = -\frac{1}{4} \sum_{i=1}^{n} \sum_{j \neq i} (y_i - y_j)^2 \tag{12}$$

and note that P^0 is just P^* when the signals are perfectly correlated.

Now suppose player i deviates to $\overline{a}_i$ in period t and thereby induces the signal distribution $\overline{q} = q\left(\cdot \mid \overline{a}_i, a^*_{-i}\right)$. Denote by $\overline{E}$ all expectations with respect to the new signal distribution $\overline{q}$ with full support over Y. Further, suppose that after the deviation, player i follows the strategy of reporting that his signal is $z_i\left(y_i\right)$ when it is actually y_i. Denoting by p^0 the normalized version of P^0 (analogous to (11)), it is easy to verify that no matter what reporting strategy z_i the deviating player follows, the expected score

$$\begin{aligned} \overline{E}\left[p^0\left(z_i\left(\widetilde{y}_i\right), \widetilde{y}_{-i}\right)\right] &= 1 + \frac{1}{2\min P_0}\left(\sum_{j\neq i}\overline{E}\left[\left(z_i\left(\widetilde{y}_i\right) - \widetilde{y}_j\right)^2\right] + \sum_{k\neq i}\sum_{j\neq k,i}\overline{E}\left[\left(\widetilde{y}_k - \widetilde{y}_j\right)^2\right]\right) \\ &< 1 \end{aligned}$$

because we have assumed that $\overline{q}$ has full support and so $\overline{E}[(z_i\left(\widetilde{y}_i\right) - \widetilde{y}_j)^2] > 0$ and $\min P^0 < 0$.

Player i's expected payoff from such a deviation is therefore

$$\begin{aligned} &(1-\delta)\, u_i\left(\overline{a}_i, a^*_{-i}\right) \\ &+\delta\Big[\overline{E}\left[p^0\left(z_i\left(\widetilde{y}_i\right), \widetilde{y}_{-i}\right)\right] u_i\left(a^*\right) + \left(1 - \overline{E}\left[p^0\left(z_i\left(\widetilde{y}_i\right), \widetilde{y}_{-i}\right)\right]\right) u_i\left(\alpha^N\right)\Big] \end{aligned} \tag{13}$$

and this is strictly smaller than $u_i\left(a^*\right)$ once δ is large enough.

Choose $\underline{\delta}_i$ such that all possible deviations $\overline{a}_i \neq a^*_i$ and all misreporting strategies z_i for i are unprofitable. Since both the actions and the signals are finite in number, such a $\underline{\delta}_i$ exists. The same is true for all players j. Let $\underline{\delta} = \max_j \underline{\delta}_j$. Then we know that for all $\delta > \underline{\delta}$, the proposed strategy profile constitutes an equilibrium. Note that $\underline{\delta}$ depends on the distributions $q\left(\cdot \mid a\right)$ for $a \neq a^*$ and not on $q\left(\cdot \mid a^*\right)$.

Nearly informative signals We can complete the proof that the given strategies form an equilibrium when q^* is nearly informative by appealing to continuity.

Recall that after renormalization, we can suppose that each $Y_i = \{1, 2, ..., K\}$. Recall that Y^D denotes the set of all diagonal profiles, that is, profiles in which all players get the same signal and that a perfectly informative distribution $q^0 \in \Delta\left(Y\right)$ is one where $q^0\left(y\right) > 0$ if and only if $y \in Y^D$. Let $Q^0 \subset \Delta\left(Y\right)$ be the set of all perfectly informative distributions.

Fix $q\left(\cdot \mid a^*\right) \equiv q^* \in \operatorname{int}\Delta\left(Y\right)$. First, find a perfectly informative distribution $q^0 \in Q^0$ such that:

$$q^0 \in \arg\min_{q\in Q^0} \|q^* - q\|_{TV}$$

and let

$$\gamma = \left\|q^* - q^0\right\|_{TV} \tag{14}$$

It is routine to verify that

$$\gamma = 1 - \sum_{y \in Y^D} q^* (y)$$

the total probability mass that q^* assigns to non-diagonal signal profiles. Moreover, $\gamma = \|q^* - q\|_{TV}$ for any $q \in Q^0$ such that for all $y \in Y^D$, $q(y) \geq q^*(y)$. Thus, in the total variation metric, q^* is equidistant from *any* perfectly informative distribution which places greater probability mass on all diagonal profiles. This means that given $q^* \in \operatorname{int} \Delta(Y)$ we can always choose a q^0 such that $q^0(y) > 0$ for all $y \in Y^D$.

Now as $\gamma \to 0$, Lemma B.1 in Appendix B guarantees that for all $y \in Y$, the score functions

$$p^*(y) \to p^0(y)$$

and so

$$E^*[p^*(\widetilde{y})] \to 1$$

as well. Moreover,

$$\max_{z_i} \overline{E}[p^*(z_i(\widetilde{y}_i), \widetilde{y}_{-i})] \to \max_{z_i} \overline{E}\left[p^0(z_i(\widetilde{y}_i), \widetilde{y}_{-i})\right]$$

Let v_i^* denote player i's payoff in the equilibrium described above. Then we have that,

$$v_i^* = (1-\delta) u_i(a^*) + \delta \Big[E^*[p^*(\widetilde{y})] v_i^* + (1 - E^*[p^*(\widetilde{y})]) u_i\left(\alpha^N\right) \Big]$$

or

$$v_i^* = \frac{(1-\delta) u_i(a^*) + \delta (1 - E^*[p^*(\widetilde{y})]) u_i\left(\alpha^N\right)}{1 - \delta E^*[p^*(\widetilde{y})]} \tag{15}$$

Thus, there exists a $\overline{\gamma}$ such that for all $\gamma < \overline{\gamma}$, for all i,

$$\begin{aligned} & (1-\delta) u_i\left(\overline{a}_i, a_{-i}^*\right) \\ & + \delta \left[\max_{z_i} \overline{E}[p^*(z_i(\widetilde{y}_i), \widetilde{y}_{-i})] v_i^* + \left(1 - \max_{z_i} \overline{E}[p^*(z_i(\widetilde{y}_i), \widetilde{y}_{-i})]\right) u_i\left(\alpha^N\right) \right] \\ < \; & (1-\delta) u_i(a^*) + \delta \Big[E^*[p^*(\widetilde{y})] v_i^* + (1 - E^*[p^*(\widetilde{y})]) u_i\left(\alpha^N\right) \Big] \end{aligned} \tag{16}$$

We have established the strategies using the score function also constitute an equilibrium thereby providing a different proof of Proposition 5.1 (under the additional condition that $E^*[\widetilde{y}_j \mid y_i]$ is increasing). Of course, the threshold values $\underline{\delta}$ and $\overline{\gamma}$ will be different than if the simple equilibrium construction were used.

Some advantages of the score function The equilibrium construction using the score function p^* overcomes some of the difficulties associated with the equilibrium using indicator function $\mathbb{I}_{Y^D}$. First, the score function is more forgiving—a slight discrepancy in the reported signals only increases the probability of punishment slightly. Second, and perhaps more important, the score function guarantees truthtelling along the path even if q^* is far from being perfectly informative (assuming, of course, that $E^*[\widetilde{y}_j \mid y_i]$ is strictly increasing). The indicator function guarantees truthtelling only when q^* is nearly informative. Specifically,

1. When there is a continuum of signals and q^* has a positive density everywhere, cooperation will break down immediately if players use the indicator function to decide on future play. This does not happen if future play is decided using the score function—now a slight disagreement in reports results in only a slight increase in the probability of future cooperation. Similarly, with a large number of players, the indicator function will result in a rapid breakdown of cooperation. Again this is not the case with the score function.

2. Recall that for the q^* defined in (8), the indicator function cannot induce truthtelling no matter how small ε is. But the score function for this example is

$$p^* = \begin{array}{c|c|c|c|} & 1 & 2 & 3 \\ \hline 1 & 1 & \frac{3-31\varepsilon}{4-20\varepsilon} & 0 \\ \hline 2 & \frac{3-31\varepsilon}{4-20\varepsilon} & \frac{1-13\varepsilon}{1-5\varepsilon} & \frac{3-31\varepsilon}{4-20\varepsilon} \\ \hline 3 & 0 & \frac{3-31\varepsilon}{4-20\varepsilon} & 1 \\ \hline \end{array}$$

   and this induces truthtelling for all $\varepsilon < \frac{1}{21}$ (this ensures that $E^*[\widetilde{y}_j \mid y_i]$ is increasing). Moreover, the conclusion of Proposition 5.1 obtains (even though q^* does not converge to a perfectly informative q^0 as ε goes to zero since $q^*(2,2) \to 0$).

6 Benefits of communication

Suppose that a^* is an efficient action and $u(a^*) \gg u(\alpha^N)$ where α^N is a Nash equilibrium of the one-shot game. The result below provides conditions under which payoffs close to $u(a^*)$ can be achieved with communication and not without. This requires the game does not have an efficient coarse correlated equilibrium.[22]

The first condition ensures that the bound developed in Proposition 4.1 is useful—that is, if the quality of monitoring is zero ($\eta = 0$), equilibrium welfare without communication are bounded away from efficient payoffs $u(a^*)$. Proposition 5.1 provides conditions under which there exists an equilibrium with communication in which players report their signals truthfully. When $q(\cdot \mid a^*)$ is perfectly informative, the equilibrium with communication results in payoffs equal to $u(a^*)$. We then have:

[22] Most games of interest satisfy the condition: the prisoners' dilemma, "chicken," Cournot oligopoly (with discrete quantities), Bertrand with or without differentiated products, etc. The rock-paper-scissors game in Example 3, on the other hand, does not satisfy the condition. In that example, $\alpha(rr) = \alpha(pp) = \alpha(ss) = \frac{1}{3}$ constitutes a coarse correlated equilibrium (but not a correlated equilibrium) that is efficient.

Theorem 6.1 *Fix $\{q(\cdot \mid a) : a \neq a^*\}$ and let $\underline{\delta}$ be defined by (13). For any $\delta > \underline{\delta}$, there exist $\overline{\eta}$ and $\overline{\gamma}$ such that if the quality of monitoring (see (2)) $\eta < \overline{\eta}$ and the informativeness of $q(\cdot \mid a^*)$ (see (14)) $\gamma < \overline{\gamma}$, then there is an equilibrium with communication whose welfare exceeds that from any equilibrium without communication.*

Note that since $q(\cdot \mid a)$ for all $a \neq a^*$ remain fixed and q^* becomes increasingly correlated as γ goes to zero, the conditions of the theorem guarantee that signals are more informative at a^* than at actions $a \neq a^*$. This feature of the signal structure is stronger than is necessary for the theorem to hold. In particular, it would be enough that q^* is "more correlated" than $q\left(\cdot \mid a_i, a_{-i}^*\right)$ for all i and all $a_i \neq a_i^*$. This is because only unilateral deviations need to be considered. But as Example 2 shows, to obtain the conclusion of the theorem, these features of the signal structure cannot be dispensed with entirely. We end with an example where the mechanics of Theorem 1 can be seen at work.

Example 4 Again consider the prisoners' dilemma

	c	d
c	$2,2$	$-1,3$
d	$3,-1$	$0,0$

but now with a monitoring structure where for both players, the set of signals is $Y_i = \{y', y'', y'''\}$ and the signal distributions are as follows:

$$q^* \equiv q(\cdot \mid cc) = \begin{array}{c|c|c|c|} & y' & y'' & y''' \\ \hline y' & \frac{1}{3}(1-\gamma) & \frac{1}{6}\gamma & \frac{1}{6}\gamma \\ \hline y'' & \frac{1}{6}\gamma & \frac{1}{3}(1-\gamma) & \frac{1}{6}\gamma \\ \hline y''' & \frac{1}{6}\gamma & \frac{1}{6}\gamma & \frac{1}{3}(1-\gamma) \\ \hline \end{array}$$

and

$$\overline{q} \equiv q(\cdot \mid \neg\, cc) = \begin{array}{c|c|c|c|} & y' & y'' & y''' \\ \hline y' & (\frac{1}{3}-\eta)^2 & \frac{1}{9}-\eta^2 & \frac{1}{9}-\frac{1}{3}\eta \\ \hline y'' & \frac{1}{9}-\eta^2 & (\frac{1}{3}+\eta)^2 & \frac{1}{9}+\frac{1}{3}\eta \\ \hline y''' & \frac{1}{9}-\frac{1}{3}\eta & \frac{1}{9}+\frac{1}{3}\eta & \frac{1}{9} \\ \hline \end{array}$$

The parameter γ is such that $\|q^* - q^0\|_{TV} = \gamma$ where q^0 is a perfectly informative distribution that places equal weight only on the diagonal elements. The parameter $\eta < \frac{1}{3}$ is such that for both i, $\|q_i^* - \overline{q}_i\|_{TV} = \eta$. Affiliation requires that $\gamma < \frac{2}{3}$.

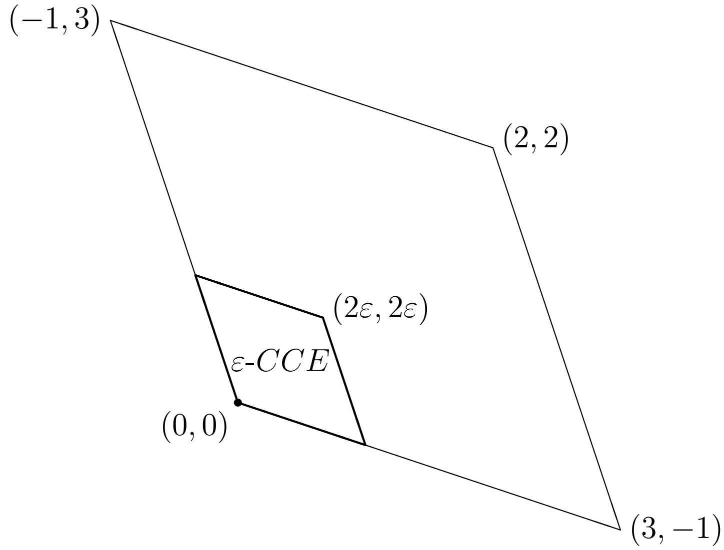

Figure 3: ε-CCE of Prisoners' Dilemma

No communication bound For the prisoners' dilemma, the set of ε-coarse correlated equilibrium payoffs is depicted in Figure 3. From Proposition 4.1 it follows that in any equilibrium of G_δ, the repeated game without communication, symmetric equilibrium payoffs $v_1 = v_2$ satisfy

$$\begin{aligned} v_i &\leq 2 \times 2\frac{\delta^2}{1-\delta}\eta \times \|u\|_\infty \\ &= 12\frac{\delta^2}{1-\delta}\eta \end{aligned}$$

since in the prisoners' dilemma, $\|u\|_\infty = 3$.

Equilibrium with communication Routine calculations show that for the monitoring structure given above, the score function $p^* : Y_1 \times Y_2 \to [0,1]$ corresponding to q^* is

$$p^* = \begin{array}{c|c|c|c|} & y' & y'' & y''' \\ \hline y' & 1 & \frac{3}{4}-\frac{3}{8}\gamma & 0 \\ \hline y'' & \frac{3}{4}-\frac{3}{8}\gamma & 1-\frac{3}{4}\gamma & \frac{3}{4}-\frac{3}{8}\gamma \\ \hline y''' & 0 & \frac{3}{4}-\frac{3}{8}\gamma & 1 \\ \hline \end{array}$$

and so that the expected score if no one deviates is $E^*\left[p^*\left(\widetilde{y}\right)\right] = 1 - \frac{3}{4}\gamma$. Using (15), we then obtain that the payoff in the equilibrium constructed in Section 5 is

$$v_i^* = \frac{2(1-\delta)}{1-\delta\left(1-\frac{3}{4}\gamma\right)}$$

Note that if instead of the score function p^*, we required that signals match in order to continue cooperating (as in Section 5.1), then the probability of continuation,

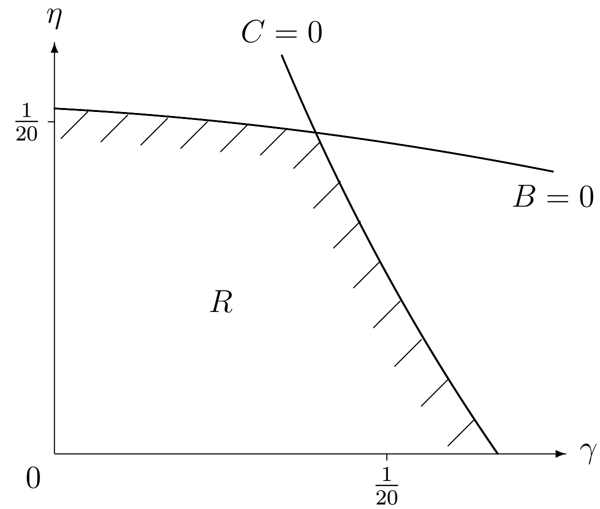

Figure 4: In Example 4, if $\delta = 0.8$, the conclusion of Theorem 1 holds when the monitoring structure is in R.

$1-\gamma$, would be lower than with the score function. The resulting equilibrium payoff would be lower as well.

If player 1, say, deviates from cc in a "normal" state, then regardless of his true signal, it is best for him to report $y_1 = y''$ in the communication phase (assuming that both γ and η are small, specifically, $4\eta + 3\gamma < 2$). This implies that, after a deviation, the probability of the state being "normal" in the next period is

$$\begin{aligned} \max_{z_1} \overline{E}\left[p^*\left(z_1\left(\widetilde{y}_1\right), \widetilde{y}_2\right)\right] &= \overline{E}\left[p^*\left(y'', \widetilde{y}_2\right)\right] \\ &= \tfrac{5}{6} + \tfrac{1}{4}\eta\left(1 - \tfrac{3}{2}\gamma\right) - \tfrac{1}{2}\gamma \end{aligned}$$

When used in (16) this implies that the suggested strategies form an equilibrium if and only if

$$\begin{aligned} B &\equiv (1-\delta)\left(u_1(cc) - u_1(dc)\right) + \delta\left[E^*\left[p^*\left(\widetilde{y}\right)\right] - \max_{z_i} \overline{E}\left[p^*\left(z_i\left(\widetilde{y}_i\right), \widetilde{y}_{-i}\right)\right]\right] v_i^* \\ &\geq 0 \end{aligned}$$

Using the expressions derived above, it is easy to verify that a necessary condition for this to hold is that $\delta \geq \frac{3}{4}$.

Payoff comparison For there to be gains from communication—equilibrium payoffs from communication exceed those from no communication—it is sufficient that

$$C \equiv v_i^* - 12\frac{\delta^2}{1-\delta}\eta \geq 0$$

For $\delta = 0.8$, Figure 4 depicts the set of parameters (γ, η) for which both $B \geq 0$ (below the $B = 0$ curve) and $C \geq 0$ (left of the $C = 0$ curve) as the shaded region labelled R. For $(\gamma, \eta) \in R$ there is an equilibrium with communication whose payoffs exceed the no-communication bound. For instance, if $\gamma = 0.02 = \eta$, equilibrium payoffs without communication are at most 0.768, whereas with communication, there is an equilibrium with payoffs of 1.887.

Our main result requires poor monitoring without communication (small η) and high correlation of signals when the actions are efficient (small γ). The roles of the two parameters are apparent in the example. A small η, of course, results in a tighter bound on the set of equilibrium payoffs without communication. A small γ has two beneficial effects when communication is possible. First, it means that the probability of staying on the equilibrium path when no one has deviated is close to one (in the example, it is $1 - \frac{3}{4}\gamma$). Second, a small γ increases the ability of players to detect deviations. A measure of this "detection ability" is $E^*\left[p^*\left(\widetilde{y}\right)\right] - \max_{z_1} \overline{E}\left[p^*\left(z_1\left(\widetilde{y}_1\right), \widetilde{y}_2\right)\right] = \frac{1}{24}\left(2 - 3\gamma\right)\left(2 - 3\eta\right)$ and this is a decreasing function of γ. This illustrates how communication improves monitoring when η and γ are small.

7 Conclusion

The general methodology of this paper is to compare the set of equilibria under two "regimes"—a base case without communication and an alternative regime in which players can communicate. This methodology can be used in a variety of applications. As already mentioned in the introduction, one may ask whether and to what extent subjective peer evaluations incentivize workers. The general methodology—especially using payoff bound developed in Section 4—can also be used to address issues other than communication. Antitrust authorities in many countries are interested in the effects of multi-market interaction among conglomerates. One may ask what the social loss from such multi-market contact is. There is a substantial literature on this subject but most of it supposes that conglomerates can perfectly monitor each other. With perfect monitoring, however, multi-market contact seems to have minimal effects.[23] How and whether multi-market contact facilitates collusion in a private monitoring set up is an open question which may be addressed by first using our CCE-bound on profits without such contact and then seeing whether there are equilibria with multi-market contact which are better for the conglomerates.

A Appendix: No-communication Bound

This appendix contains the proof of Lemma 4.1 from Section 4 which provides an estimate of the "loss from punishment." In other words, it measures the payoff difference to a permanently deviating player from facing the actual strategy versus the fictitious "non-responsive" strategy. Each of these generate different distributions over histories and we begin by estimating the distance between these distributions. Lemma A.1

[23] Matsushima (2001) analyzes multi-market contact with public monitoring.

shows that the distance between these distributions grows at most linearly. Lemma A.2 then derives the bound in Lemma 4.1.

For a fixed strategy profile σ, let $\mu^t(h^t)$ be the probability that a history h^t is realized. Then

$$\mu^t\left(h^t\right) = \mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t \mid h^{t-1}\right) q\left(y^t \mid a^t\right)$$

where $h^t = (h^{t-1}, a^t, y^t)$. Let

$$\begin{aligned}\mu^t_{-i}\left(h^t_{-i}\right) &= \sum_{h^t_i} \mu^t\left(h^t\right) \\ &= \sum_{h^{t-1}_i}\sum_{a^t_i} \mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t \mid h^{t-1}\right) q_{-i}\left(y^t_{-i} \mid a^t\right)\end{aligned}$$

be the marginal distribution of μ^t on the private histories $h^t_{-i} = \left(h^{t-1}_{-i}, a^t_{-i}, y^t_{-i}\right)$ of players $j \neq i$.

Similarly, let $\overline{\mu}^t$ be the probability of history h^t that results when i permanently deviates to $\overline{a}_i$, that is, from the strategy profile $(\overline{\sigma}_i, \sigma_{-i})$. Then,

$$\overline{\mu}^t\left(h^t\right) = \begin{cases} \overline{\mu}^{t-1}\left(h^{t-1}\right)\sigma_{-i}\left(a^t_{-i} \mid h^{t-1}_{-i}\right) q\left(y^t \mid a^t\right) & \text{if } a^t_i = \overline{a}_i \\ 0 & \text{otherwise}\end{cases}$$

and analogously let $\overline{\mu}^t_{-i}$ be the marginal distribution of $\overline{\mu}^t$ on player $-i$'s private histories h^t_{-i} so that

$$\begin{aligned}\overline{\mu}^t_{-i}\left(h^t_{-i}\right) &= \sum_{h^t_i} \overline{\mu}^t\left(h^t\right) \\ &= \overline{\mu}^{t-1}_{-i}\left(h^{t-1}_{-i}\right)\sigma_{-i}\left(a^t_{-i} \mid h^{t-1}_{-i}\right) q_{-i}\left(y^t_{-i} \mid \overline{a}_i, a^t_{-i}\right)\end{aligned}$$

Lemma A.1 *For all t, $\left\|\mu^t_{-i} - \overline{\mu}^t_{-i}\right\|_{TV} \leq t\eta$.*

Proof. The proof is by induction on t. For $t = 1$, we have

$$\begin{aligned}\mu^1_{-i}\left(h^1_{-i}\right) - \overline{\mu}^1_{-i}\left(h^1_{-i}\right) &= \sum_{a^1_i} \sigma\left(a^1\right) q_{-i}\left(y^1_{-i} \mid a^1\right) - \sigma_{-i}\left(a^1_{-i}\right) q_{-i}\left(y^1_{-i} \mid \overline{a}_i, a^1_{-i}\right) \\ &= \sum_{a^1_i} \sigma\left(a^1\right) q_{-i}\left(y^1_{-i} \mid a^1\right) - \sum_{a^1_i} \sigma\left(a^1\right) q_{-i}\left(y^1_{-i} \mid \overline{a}_i, a^1_{-i}\right) \\ &= \sum_{a^1_i} \sigma\left(a^1\right)\left(q_{-i}\left(y^1_{-i} \mid a^1\right) - q_{-i}\left(y^1_{-i} \mid \overline{a}_i, a^1_{-i}\right)\right)\end{aligned}$$

since $\sigma_{-i}\left(a_{-i}^1\right)=\sum_{a_i^1}\sigma\left(a^1\right)$ and so

$$\begin{aligned}
\left\|\mu_{-i}^1-\overline{\mu}_{-i}^1\right\|_{TV} &= \frac{1}{2}\sum_{h_{-i}^1}\left|\mu_{-i}^1\left(h_{-i}^1\right)-\overline{\mu}_{-i}^1\left(h_{-i}^1\right)\right| \\
&\leq \frac{1}{2}\sum_{h_{-i}^1}\sum_{a_i^1}\sigma\left(a^1\right)\left|q_{-i}\left(y_{-i}^1\mid a^1\right)-q_{-i}\left(y_{-i}^1\mid \overline{a}_i,a_{-i}^1\right)\right| \\
&= \frac{1}{2}\sum_{a_{-i}^1}\sum_{a_i^1}\sigma\left(a^1\right)\sum_{y_{-i}^1}\left|q_{-i}\left(y_{-i}^1\mid a^1\right)-q_{-i}\left(y_{-i}^1\mid \overline{a}_i,a_{-i}^1\right)\right| \\
&= \sum_{a^1}\sigma\left(a^1\right)\left\|q_{-i}\left(\cdot\mid a^1\right)-q_{-i}\left(\cdot\mid \overline{a}_i,a_{-i}^1\right)\right\|_{TV} \\
&\leq \eta
\end{aligned}$$

the last inequality follows from our assumption that the quality of the monitoring does not exceed η.

Now suppose that the statement of the lemma holds for $t-1$. We can write

$$\begin{aligned}
\Delta = {} & \mu_{-i}^t\left(h_{-i}^t\right)-\overline{\mu}_{-i}^t\left(h_{-i}^t\right) \\
= {} & \sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t\mid h^{t-1}\right)q_{-i}\left(y_{-i}^t\mid a^t\right) \\
& -\overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\sigma_{-i}\left(a_{-i}^t\mid h_{-i}^{t-1}\right)q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right) \\
= {} & \sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t\mid h^{t-1}\right)q_{-i}\left(y_{-i}^t\mid a^t\right) \\
& -\sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t\mid h^{t-1}\right)q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right) \\
& +\sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t\mid h^{t-1}\right)q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right) \\
& -\overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\sigma_{-i}\left(a_{-i}^t\mid h_{-i}^{t-1}\right)q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right)
\end{aligned}$$

Combining terms, we have

$$\begin{aligned}
\Delta &= \sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t\mid h^{t-1}\right)\left[q_{-i}\left(y_{-i}^t\mid a^t\right)-q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right)\right] \\
&\quad+\left[\sum_{h_i^{t-1}}\mu^{t-1}\left(h^{t-1}\right)\sigma_{-i}\left(a_{-i}^t\mid h_{-i}^{t-1}\right)-\overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\sigma_{-i}\left(a_{-i}^t\mid h_{-i}^{t-1}\right)\right]q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right) \\
&= \sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t\mid h^{t-1}\right)\left[q_{-i}\left(y_{-i}^t\mid a^t\right)-q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right)\right] \\
&\quad+\left[\mu_{-i}^{t-1}\left(h_{-i}^{t-1}\right)-\overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\right]\sigma_{-i}\left(a_{-i}^t\mid h_{-i}^{t-1}\right)q_{-i}\left(y_{-i}^t\mid \overline{a}_i,a_{-i}^t\right)
\end{aligned}$$

where we have used the fact that $\sum_{a_i^t} \sigma\left(a^t \mid h^{t-1}\right) = \sum_{a_i^t} \sigma_i\left(a_i^t \mid h_i^{t-1}\right) \sigma_{-i}\left(a_{-i}^t \mid h_{-i}^{t-1}\right) = \sigma_{-i}\left(a_{-i}^t \mid h_{-i}^{t-1}\right).$

Thus,

$$\begin{aligned}
& \left\|\mu_{-i}^t - \overline{\mu}_{-i}^t\right\|_{TV} \\
= \ & \frac{1}{2}\sum_{h_{-i}^t}\left|\mu_{-i}^t\left(h_{-i}^t\right) - \overline{\mu}_{-i}^t\left(h_{-i}^t\right)\right| \\
\leq \ & \frac{1}{2}\sum_{h_{-i}^t}\sum_{h_i^{t-1}}\sum_{a_i^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t \mid h^{t-1}\right)\left|q_{-i}\left(y_{-i}^t \mid a^t\right) - q_{-i}\left(y_{-i}^t \mid \overline{a}_i, a_{-i}^t\right)\right| \\
& + \frac{1}{2}\sum_{h_{-i}^t}\left|\mu_{-i}^{t-1}\left(h_{-i}^{t-1}\right) - \overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\right|\sigma_{-i}\left(a_{-i}^t \mid h_{-i}^{t-1}\right) q_{-i}\left(y_{-i}^t \mid \overline{a}_i, a_{-i}^t\right)
\end{aligned}$$

Since $h_{-i}^t = \left(h_{-i}^{t-1}, a_{-i}^t, y_{-i}^t\right)$, the first term equals

$$\begin{aligned}
& \sum_{h^{t-1}}\sum_{a^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t \mid h^{t-1}\right)\left\|q_{-i}\left(\cdot \mid a^t\right) - q_{-i}\left(\cdot \mid \overline{a}_i, a_{-i}^t\right)\right\|_{TV} \\
\leq \ & \sum_{h^{t-1}}\sum_{a^t}\mu^{t-1}\left(h^{t-1}\right)\sigma\left(a^t \mid h^{t-1}\right)\eta \\
= \ & \eta
\end{aligned}$$

The second term equals

$$\begin{aligned}
& \frac{1}{2}\sum_{h_{-i}^{t-1}}\sum_{a_{-i}^t}\sum_{y_{-i}^t}\left|\mu_{-i}^{t-1}\left(h_{-i}^{t-1}\right) - \overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\right|\sigma_{-i}\left(a_{-i}^t \mid h_{-i}^{t-1}\right) q_{-i}\left(y_{-i}^t \mid \overline{a}_i, a_{-i}^t\right) \\
= \ & \frac{1}{2}\sum_{h_{-i}^{t-1}}\left|\mu_{-i}^{t-1}\left(h_{-i}^{t-1}\right) - \overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\right|\sum_{a_{-i}^t}\sum_{y_{-i}^t}\sigma_{-i}\left(a_{-i}^t \mid h_{-i}^{t-1}\right) q_{-i}\left(y_{-i}^t \mid \overline{a}_i, a_{-i}^t\right) \\
= \ & \frac{1}{2}\sum_{h_{-i}^{t-1}}\left|\mu_{-i}^{t-1}\left(h_{-i}^{t-1}\right) - \overline{\mu}_{-i}^{t-1}\left(h_{-i}^{t-1}\right)\right| \\
= \ & \left\|\mu_{-i}^{t-1} - \overline{\mu}_{-i}^{t-1}\right\|_{TV} \\
\leq \ & (t-1)\,\eta
\end{aligned}$$

This completes the proof. ■

Lemma 4.1 is restated as

Lemma A.2 *Suppose i plays $\overline{a}_i$ always. The difference in i's payoff when others play σ_{-i} versus when they play the non-responsive strategy α_{-i} derived from σ satisfies*

$$\left|v_i\left(\overline{\sigma}_i, \sigma_{-i}\right) - v_i\left(\overline{\sigma}_i, \alpha_{-i}\right)\right| \leq 2\frac{\delta^2}{1-\delta}\eta \times \|u\|_\infty$$

Proof.

$$\begin{aligned}
v_i(\overline{\sigma}_i, \sigma_{-i}) &= (1-\delta)\sum_{t=1}^{\infty}\sum_{h_{-i}^{t-1}}\sum_{a_{-i}}\delta^t u_i(\overline{a}_i, a_{-i})\,\sigma_{-i}\left(a_{-i} \mid h_{-i}^{t-1}\right)\overline{\mu}_{-i}\left(h_{-i}^{t-1}\right) \\
&= (1-\delta)\sum_{t=1}^{\infty}\delta^t\sum_{h_{-i}^{t-1}}\left(\sum_{a_{-i}}u_i(\overline{a}_i, a_{-i})\,\sigma_{-i}\left(a_{-i} \mid h_{-i}^{t-1}\right)\right)\overline{\mu}_{-i}\left(h_{-i}^{t-1}\right) \\
&= (1-\delta)\sum_{t=1}^{\infty}\delta^t\sum_{h_{-i}^{t-1}}E\left[u_i(\overline{a}_i, \sigma_{-i}) \mid h_{-i}^{t-1}\right]\overline{\mu}_{-i}\left(h_{-i}^{t-1}\right)
\end{aligned}$$

Similarly,

$$v_i(\overline{\sigma}_i, \alpha_{-i}) = (1-\delta)\sum_{t=1}^{\infty}\sum_{a_{-i}}\delta^t u_i(\overline{a}_i, a_{-i})\,\alpha_{-i}^t(a_{-i})$$

and since

$$\alpha_{-i}^t(a_{-i}) = \sum_{h_{-i}^{t-1}}\sigma_{-i}\left(a_{-i} \mid h_{-i}^{t-1}\right)\mu_{-i}\left(h_{-i}^{t-1}\right)$$

we have

$$\begin{aligned}
v_i(\overline{\sigma}_i, \alpha_{-i}) &= (1-\delta)\sum_{t=1}^{\infty}\sum_{h_{-i}^{t-1}}\sum_{a_{-i}}\delta^t u_i(\overline{a}_i, a_{-i})\,\sigma_{-i}\left(a_{-i} \mid h_{-i}^{t-1}\right)\mu_{-i}\left(h_{-i}^{t-1}\right) \\
&= (1-\delta)\sum_{t=1}^{\infty}\delta^t\sum_{h_{-i}^{t-1}}\left(\sum_{a_{-i}}u_i(\overline{a}_i, a_{-i})\,\sigma_{-i}\left(a_{-i} \mid h_{-i}^{t-1}\right)\right)\mu_{-i}\left(h_{-i}^{t-1}\right) \\
&= (1-\delta)\sum_{t=1}^{\infty}\delta^t\sum_{h_{-i}^{t-1}}E\left[u_i(\overline{a}_i, \sigma_{-i}) \mid h_{-i}^{t-1}\right]\mu_{-i}\left(h_{-i}^{t-1}\right)
\end{aligned}$$

and so

$$\begin{aligned}
&\left|v_i(\overline{\sigma}_i, \sigma_{-i}) - v_i(\overline{\sigma}_i, \alpha_{-i})\right| \\
\leq\;& (1-\delta)\sum_{t=1}^{\infty}\delta^t\left|\sum_{h_{-i}^{t-1}}E\left[u_i(\overline{a}_i, \sigma_{-i}) \mid h_{-i}^{t-1}\right]\left(\overline{\mu}_{-i}\left(h_{-i}^{t-1}\right) - \mu_{-i}\left(h_{-i}^{t-1}\right)\right)\right| \\
\leq\;& 2(1-\delta)\sum_{t=1}^{\infty}\delta^t\left\|\mu_{-i}^{t-1} - \overline{\mu}_{-i}^{t-1}\right\|_{TV} \times \|u\|_{\infty}
\end{aligned}$$

where the second inequality follows from the fact that given any two measures μ and $\overline{\mu}$, $|E_{\mu}[f] - E_{\overline{\mu}}[f]| \leq 2\|\mu - \overline{\mu}\|_{TV} \times \|f\|_{\infty}$ for any f such that the expectations are well-defined (see, for instance, Levin et al., 2009).[24]

[24] The property that a linear function of the TV distance is an upper bound to the difference in payoffs account for the fact that the bound in the statement of Lemma A.2 is linear in this distance.

Now Lemma A.1 shows that $\left\|\mu_{-i}^{t-1}-\overline{\mu}_{-i}^{t-1}\right\|_{TV}\leq(t-1)\,\eta$. Thus, we obtain[25]

$$\begin{aligned}\left|v_i\left(\overline{\sigma}_i,\sigma_{-i}\right)-v_i\left(\overline{\sigma}_i,\alpha_{-i}\right)\right| &\leq 2\left(1-\delta\right)\sum_{t=1}^{\infty}\delta^t\left(t-1\right)\eta\left\|u\right\|_\infty \\ &= 2\frac{\delta^2}{1-\delta}\eta\times\left\|u\right\|_\infty\end{aligned}$$

■

B Appendix: Communication Equilibrium

In this appendix, we first show that the score function defined in (10) has the property that every player has the strict incentive to report his private signal truthfully at the communication stage, provided that the player chose a_i^*, his part of the efficient action a^* (see Proposition B.1).[26]

A second result, Lemma B.1 below, establishes a continuity result used in the proof of Proposition 5.1. It shows that the score functions associated with distributions that are nearly informative are close to score functions of perfectly informative distributions.

Let $Y_i=\{1,2,...,K\}$. Given a distribution $q\in\Delta(Y)$, consider the score function (as in (10))

$$P\left(y\right)=\frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i}y_iy_j-\sum_{i=1}^{n}\sum_{x_i<y_i}\phi_i\left(x_i\right)-\frac{1}{2}\sum_{i=1}^{n}\phi_i\left(y_i\right)\tag{17}$$

Proposition B.1 *Suppose $q\in\operatorname{int}\Delta(Y)$ satisfies the condition that $E\left[\widetilde{y}_j\mid y_i\right]$ is increasing in y_i. Then the score function defined in (17) induces strict truth-telling.*

Proof. Given q, define $\phi_i:Y_i\to\mathbb{R}$ by

$$\phi_i\left(y_i\right)=E\left[\textstyle\sum_{j\neq i}\widetilde{y}_j\mid y_i\right]$$

as the expectation (with respect to q) of the sum of the other players' choices, conditional on $\widetilde{y}_i=y_i$.

[25] The bound below is, of course, an overestimate since the TV distance between two measures is always less than or equal to 1.

[26] The fact that all players have the same "score" distinguishes this result from that of Crémer and McLean (1988).

Suppose all players $j \neq i$ report truthfully. When player i receives a signal y_i and reports y_i', his expected score is[27]

$$\begin{aligned} E\left[P\left(y_i', \widetilde{y}_{-i}\right) \mid \widetilde{y}_i = y_i\right] &= y_i' E\left[\sum_{j \neq i} \widetilde{y}_j \mid y_i\right] - \sum_{x_i < y_i'} \phi_i\left(x_i\right) - \frac{1}{2}\phi_i\left(y_i'\right) - C_i\left(y_i\right) \\ &= y_i'\phi_i\left(y_i\right) - \sum_{x < y_i'} \phi_i\left(x_i\right) - \frac{1}{2}\phi_i\left(y_i'\right) - C_i\left(y_i\right) \end{aligned}$$

where $C_i\left(y_i\right)$ are the terms that depend on the conditional distribution $q\left(\cdot \mid y_i\right) \in \Delta\left(Y_{-i}\right)$ but do *not* depend on player i's action y_i'. On the other hand, if after receiving a signal y_i, player i tells the truth and reports y_i, his expected score is

$$E\left[P\left(y_i, \widetilde{y}_{-i}\right) \mid y_i\right] = y_i\phi_i\left(y_i\right) - \sum_{x < y_i} \phi_i\left(x\right) - \frac{1}{2}\phi_i\left(y_i\right) - C_i\left(y_i\right)$$

Let

$$\Delta = E\left[P\left(y_i, \widetilde{y}_{-i}\right) \mid y_i\right] - E\left[P\left(y_i', \widetilde{y}_{-i}\right) \mid y_i\right]$$

First, suppose $y_i' < y_i$. Then,

$$\begin{aligned} \Delta &= \left(y_i - y_i'\right)\phi_i\left(y_i\right) - \sum_{y_i' \leq x < y_i} \phi_i\left(x\right) - \frac{1}{2}\left(\phi_i\left(y_i\right) - \phi_i\left(y_i'\right)\right) \\ &= \sum_{y_i' \leq x < y_i} \left(\phi_i\left(y_i\right) - \phi_i\left(x\right)\right) - \frac{1}{2}\left(\phi_i\left(y_i\right) - \phi_i\left(y_i'\right)\right) \\ &= \sum_{y_i' < x < y_i} \left(\phi_i\left(y_i\right) - \phi_i\left(x\right)\right) + \frac{1}{2}\left(\phi_i\left(y_i\right) - \phi_i\left(y_i'\right)\right) \\ &> 0 \end{aligned}$$

since ϕ_i is strictly increasing.

Now suppose $y_i' > y_i$. Then

$$\begin{aligned} \Delta &= -\left(y_i' - y_i\right)\phi_i\left(y_i\right) + \sum_{y_i < x \leq y_i'} \phi_i\left(x\right) + \frac{1}{2}\left(\phi_i\left(y_i'\right) - \phi_i\left(y_i\right)\right) \\ &= \sum_{y_i < x \leq y_i'} \left(\phi_i\left(x\right) - \phi_i\left(y_i\right)\right) + \frac{1}{2}\left(\phi_i\left(y_i'\right) - \phi_i\left(y_i\right)\right) \\ &> 0 \end{aligned}$$

again since ϕ_i is strictly increasing.

Thus, upon receiving the signal y_i, player i has a strict incentive to report truthfully. ■

[27] Note that $\frac{1}{2}\sum_{i \neq j} y_i y_j = y_i \sum_{j \neq i} y_j + \frac{1}{2}\sum_{j \neq k} y_j y_k$ where in the last sum $j \neq i$ and $k \neq i$.

The next result establishes the continuity of score functions in the underlying distributions.

Lemma B.1 *Let $\left\{q^l\right\} \in \operatorname{int}\Delta(Y)$ be a sequence such that $q^l \to q^0$ and $q^0 \in \Delta(Y)$ is a perfectly informative distribution. Then the corresponding score functions $p^l \to p^0$.*

Proof. Analogous to (10), we have that

$$P^l(y) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i} y_i y_j - \sum_{i=1}^{n}\sum_{x_i<y_i} E^l\left[\sum\nolimits_{j\neq i}\widetilde{y}_j \mid x_i\right] - \frac{1}{2}\sum_{i=1}^{n} E^l\left[\sum\nolimits_{j\neq i}\widetilde{y}_j \mid y_i\right] \quad (18)$$

where E^l denotes expectations under q^l. On the other hand, notice that

$$P^0(y) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i} y_i y_j - \sum_{i=1}^{n}\sum_{x_i<y_i}(n-1)x_i - \frac{1}{2}\sum_{i=1}^{n}(n-1)y_i \quad (19)$$

because the right-hand side equals

$$\begin{aligned} & \frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i} y_i y_j - (n-1)\sum_{i=1}^{n}\frac{(y_i-1)y_i}{2} - \frac{1}{2}(n-1)\sum_{i=1}^{n} y_i \\ = & \frac{1}{2}\sum_{i=1}^{n}\sum_{j\neq i} y_i y_j - \frac{1}{2}(n-1)\sum_{i=1}^{n} y_i^2 \\ = & -\frac{1}{4}\sum_{i=1}^{n}\sum_{j\neq i}(y_i-y_j)^2 \end{aligned}$$

Comparing terms in (18) and (19) shows that as $q^l \to q^0$, $E^l[\widetilde{y}_j \mid \widetilde{y}_i = x_i] \to x_i$ (recall that each x_i occurs with positive probability).

Finally, note that since $\max P^0 - \min P^0 > 0$, $\lim\left(\max P^l - \min P^l\right) > 0$, and $\lim p^l$ is well defined. Thus, p^l (the normalized version of P^l) converges to p^0 (the normalized version of P^0) as well. ■

References

[1] Aoyagi, Masaki (2002): "Collusion in Dynamic Bertrand Oligopoly with Correlated Private Signals and Communication," *Journal of Economic Theory*, 102, 229–248.

[2] Aryal, Gaurab, Federico Ciliberto and Benjamin Leyden (2018): "Public Communication and Collusion in the Airline Industry," SSRN Working Paper 3122560.

[3] Aumann, Robert (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55 (1), 1–18.

[4] Aumann, Robert and Michael Maschler (1995): *Repeated Games with Incomplete Information*, MIT Press.

[5] Awaya, Yu (2014): "Private Monitoring and Communication in Repeated Prisoners' Dilemma." https://sites.google.com/site/yuawaya/research

[6] Awaya, Yu and Vijay Krishna (2016): "On Communication and Collusion," *American Economic Review*, 106 (2), 285–315.

[7] Blackwell, David (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, University of California Press, 93–102.

[8] Compte, Olivier (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66 (3), 597–626.

[9] Crémer, Jacques, and Richard McLean (1988): "Full Extraction of the Surplus in Bayesian and Dominant Strategy Auctions," *Econometrica*, 56 (6): 1247–1257.

[10] Deb, Joyee, Jin Li and Arijit Mukherjee (2016): "Relational Contracts with Subjective Peer Evaluations," *Rand Journal of Economics*, 47 (1), 3–28.

[11] Fleckinger, Pierre (2012): "Correlation and Relative Performance Evaluation," *Journal of Economic Theory*, 147, 93–117.

[12] Fudenberg, Drew and David Levine (2007): "The Nash-Threats Folk Theorem with Communication and Approximate Common Knowledge in Two Player Games," *Journal of Economic Theory*, 132, 461–473.

[13] Gérard-Varet, Louis-André and Hervé Moulin (1978): "Correlation in Duopoly," *Journal of Economic Theory*, 19 (1), 123–139.

[14] Gromb, Denis and David Martimort (2007): "Collusion and the Organization of Delegated Expertise," *Journal of Economic Theory*, 137, 271–299.

[15] Hart, Sergiu and David Schmeidler (1989): "Existence of Correlated Equilibria," *Mathematics of Operations Research*, 14 (1), 18–25.

[16] Hörner, Johannes and Wojciech Olzewski (2009): "How Robust is the Folk Theorem?" *Quarterly Journal of Economics*, 124 (4), 1773–1814.

[17] Kandori, Michihiro and Hitoshi Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66 (3), 627–652.

[18] Kandori, Michihiro (2003): "Randomization, Communication and Efficiency in Repeated Games with Imperfect Public Monitoring," *Econometrica*, 71 (1), 345–353.

[19] Levin, David, Yuval Peres and Elizabeth Wilmer (2009): *Markov Chains and Mixing Times*, American Mathematical Society.

[20] Mailath, George and Stephen Morris (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102 (1), 189–228.

[21] Mailath, George and Larry Samuelson (2006): *Repeated Games and Reputations*, Oxford University Press.

[22] Matsushima, Hitoshi (2001): "Multimarket Contact, Imperfect Monitoring and Implicit Collusion," *Journal of Economic Theory*, 98, 158–178.

[23] Monderer, Dov and Lloyd Shapley (1996): "Potential Games," *Games and Economic Behavior*, 14, 124–143.

[24] Moulin, Hervé and Jean-Philippe. Vial (1978): "Strategically Zero-Sum Games: The Class of Games Whose Mixed Strategy Equilibria Cannot be Improved Upon," *International Journal of Game Theory*, 7 (3/4), 201–221.

[25] Moulin, Hervé, Indrajit Ray and Sonali Sen Gupta (2014): "Improving Nash by Coarse Correlation," *Journal of Economic Theory*, 150, 852–865.

[26] Neyman, Abraham (1997): "Correlated Equilibrium and Potential Games," *International Journal of Game Theory*, 26 (2), 223–227.

[27] Obara, Ichiro (2009): "Folk Theorem with Communication," *Journal of Economic Theory*, 144, 120–134.

[28] Ostrom, Elinor, Roy Gardner and James Walker (1994): *Rules, Games and Common Pool Resources*, University of Michigan Press.

[29] Pai, Mallesh, Aaron Roth and Jonathan Ullman (2016): "An Anti-Folk Theorem for Large Repeated Games with Imperfect Monitoring," *ACM Transactions in Economics and Computation*, 5 (2).

[30] Rahman, David (2014): "The Power of Communication," *American Economic Review*, 104 (11), 3737–3751.

[31] Roughgarden, Tim (2016): *Twenty Lectures on Algorithmic Game Theory*, Cambridge University Press.

[32] Sekiguchi, Tadashi (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76, 345–361.

[33] Sion, Maurice (1958): "On General Minmax Theorems," *Pacific Journal of Mathematics*, 8 (1), 171–176.

[34] Spector, David (2015): "Cheap talk, Monitoring and Collusion," Working Paper, Paris School of Economics.

[35] Stigler, George (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 74, 44–61.

[36] Sugaya, Takuo (2015): "Folk Theorem in Repeated Games with Private Monitoring," Stanford University, https://sites.google.com/site/takuosugaya/home/research

[37] Sugaya, Takuo and Alexander Wolitzky (2017): " Bounding Equilibrium Payoffs in Repeated Games with Private Monitoring," *Theoretical Economics*, 12 (2), 691–729.

[38] Zheng, Bingyong (2008): "Approximate Efficiency in Repeated Games with Correlated Private Signals," *Games and Economic Behavior*, 63, 406–416.