Theoretical Economics 11 (2016), 53–87 1555-7561/20160053

Approximate efficiency in repeated games with side-payments and correlated signals

Jimmy H. Chan
School of Economics, Fudan University

Wenzhang Zhang
School of Economics, Shanghai University of Finance and Economics

Side-payments are common in many long-term relationships. We show that when players can exchange side-payments, approximate efficiency is achievable in any repeated game with private monitoring and communication, as long as the players can observe their own payoffs and are sufficiently patient, the efficient stage-game outcome is unique, and the signal distribution has full support. Unlike existing results in the literature, our result does not require deviations to be statistically detectable.

Keywords. Repeated games, private monitoring, communication.

JEL classification. C73.

1. Introduction

Motivated by the secret-price-cuts problem (Stigler 1964), researchers have studied repeated games with private monitoring and communication. A standard approach in the literature is to require unilateral deviations to be statistically detectable. In this paper, we introduce an alternative approach that exploits information embedded in payoff functions. We show that approximate efficiency can be achieved if the following conditions are met:

- Each player can observe his own stage-game payoff.
- Any unilateral deviation from the efficient outcome strictly lowers the total payoff.
- Players can exchange side-payments at the end of each period.
- The signal distribution has full support.

The first and second conditions are natural in many economic problems. While side-payments between players are untypical in repeated games, they are in fact quite common in reality. For example, employers pay workers monetary bonuses in relational

Jimmy H. Chan: jimmy.hing.chan@gmail.com
Wenzhang Zhang: wenzhang.zhang@gmail.com
We thank the co-editor, three anonymous referees, and participants at various seminars and conferences for helpful comments. This research is funded by the Chinese National Science Foundation (Project 71171125). Chan is supported by the Shanghai Dongfang Xuezhe Program.

DOI: 10.3982/TE1369

contracts (Levin 2003 and Fuchs 2007) and cartels exchange side-payments through interfirm trades (Harrington 2006, Harrington and Skrzypacz 2007, and Harrington and Skrzypacz 2011). The last condition, which is standard in the literature, ensures that no player can use his own signal to rule out any signal of another player. Unlike existing results in the literature, our result does not assume statistical detectability.

To understand our result, notice that statistical detectability is not needed to enforce an efficient outcome; we can achieve that by simply setting the continuation payoff of each player equal to the total payoff of the other players (Fudenberg et al. 1994 and Chan and Zhang 2012). The key issue is how to mitigate the efficiency loss caused by imperfect monitoring. One approach is to use a T-period delayed-communication mechanism to link decisions across periods (Compte 1998, Obara 2009, and Fong et al. 2011). In this paper, we extend this approach to allow for any correlation structure of the private signals.

The mechanism consists of three components. The first is a delayed-communication mechanism similar to that in Fong et al. (2011). It requires that when any player j reports a discounted payoff that is below a target at the end of period T, each player $i \neq j$ must pay a penalty equal to the shortfall. Each player i's transfer is thus linear in the discounted payoff of each player $j \neq i$ up to player j's target. The targets are set slightly above the equilibrium expected discounted payoffs so that it is extremely unlikely for the discounted payoff of any player to exceed his target when T is large. But while unlikely ex ante, during the mechanism a player may still learn through his own signals that some player's payoff is likely to be above target. When that happens, the player may want to deviate from the efficient action.

The second component of our mechanism, the main contribution of this paper, is designed to deal with this "learning" problem. It can be loosely described as "side bets" between each pair of players. The bets have the following properties. In any period and for each player pair ij, whenever player i believes player j's total discounted payoff is likely to be above target, he also believes he will almost surely receive from player j a bonus equal to the amount of player j's payoff that is above target; however, regardless of the signals he has observed, player j always believes that the probability of paying the bonus is constant and equal to the ex ante probability of paying the bonus. Since the probability is very low when T is large, player j has little incentive to deviate or misreport to avoid paying player i. But player i now expects that his transfer is almost always equal to the total payoff of the other players (minus a constant). Since the efficient outcome is strict, it is a best response for i to choose the efficient action in every period, regardless of whether he is going to lie about his signals. Finally, the third component of our mechanism is a scoring rule that induces the players to report their signals truthfully.[1]

We establish our approximate-efficiency result by implementing the T-period mechanism in a repeated game with private monitoring and communication. Since the players' equilibrium discounted payoffs are going to be very close to the mean, they are unlikely to be punished by the first component. Although the second component is not fully efficient—the amount player j has to pay is actually greater than what player i

[1] Nissim et al. (2012) develop a similar approach for approximate implementation in large societies.

receives—the efficiency loss is small because player j rarely has to pay player i in equilibrium. The efficiency loss due to the scoring rule is also very small because the incentive to misreport is very weak.

A limitation of our mechanism is that it requires negative transfers. In standard repeated games without side-payments, one may need to enforce an inefficient outcome with positive transfers to redistribute payoffs. The availability of side-payments allows us to redistribute payoffs directly and avoid positive transfers. Without side-payments, our result may still hold under additional restrictions on the feasible payoff set. We consider the no-side-payment case in Section 5.

1.1 *Related literature*

Our mechanism is similar to the classic Clarke–Groves mechanism except that it deals with hidden actions and not private information. Bergemann and Välimäki (2010) and Athey and Segal (2013) extend the Clarke–Groves mechanism to a dynamic environment, but their models involve only private information. An early version of Athey and Segal (2013) allows the agents to take private actions. Their results do not apply to our model because in their model, the action of one agent does not directly affect the payoffs of the other agents.

The idea of reducing efficiency loss by linking decisions across periods was first introduced by Rubinstein and Yaari (1983) and Radner (1985) in one-sided repeated moral-hazard problems, and by Abreu et al. (1991) in repeated games with imperfect monitoring.[2] Compte (1998) was the first to apply this idea to repeated games with private monitoring and communication. Since Compte (1998) assumes the players' signals are independent, there is no learning within a T-period block. One way to deal with the learning problem is to punish the "optimistic" players—those who think they are unlikely to be punished—more heavily. However, if an optimistic player is punished more heavily when he honestly reveals his belief, he would have incentives to lie. Zheng (2008) provides a sufficient condition on the monitoring structure in symmetric games under which this type of lying can be deterred. Obara (2009) provides a more general condition, which requires that, for at least one player, deviations and false reports be jointly statistically detectable. Our "side-bets" approach, by contrast, solves the learning problem by exploiting the correlation that causes the problem in the first place.

Our result extends the technique of reducing efficiency loss through delayed communication. An alternative approach to reduce efficiency loss is due to Kandori and Matsushima (1998) and Fudenberg et al. (1994).[3] Other related works with

[2] In repeated moral-hazard problems, linking decisions improves efficiency by allowing the mechanism designer to use the same punishment to motivate the players in multiple periods. In repeated private-information problems, linking decisions is also useful but through a different channel. See Jackson and Sonnenschein (2007) and Escobar and Toikka (2013).

[3] They show that if deviations are pairwise identifiable, it is possible to transfer the punishment of one player to another so that the total punishment is always zero. Hörner et al. (2013) extend this approach to incorporate Markovian private information.

communication include Ben-Porath and Kahneman (1996), Aoyagi (2002), and Fudenberg and Levine (2007). Efficiency results and folk theorems have also been proven without communication. Earlier works have focused on the cases where signals are almost perfect (Sekiguchi 1997, Bhaskar and van Damme 2002, Bhaskar and Obara 2002, Ely and Välimäki 2002, Piccione 2002, Hörner and Olszewski 2006, and Yamamoto 2007), almost public (Mailath and Morris 2002 and Hörner and Olszewski 2009), or independent (Matsushima 2004 and Yamamoto 2007). See Kandori (2002) and Mailath and Samuelson (2006) for excellent surveys of this literature. Two recent advances allow for correlated signals. Fong et al. (2011) show that approximate efficiency can be attained in the repeated two-player prisoners' dilemma game when the monitoring structure is sufficiently accurate. Sugaya (2012) proves a folk theorem without communication for general finite-player repeated games under a set of conditions that includes statistical detectability. Since we do not assume statistical detectability, the result of Sugaya (2012) does not directly apply. It is an open question whether approximate efficiency is achievable in our model without communication.

2. The model

A group of n players, denoted by $\mathcal{N} = \{1, 2, \dots, n\}$, play the following stage game G in each period $t = 1, 2, \dots$. First, each player $i \in \mathcal{N}$ simultaneously chooses a private action a_i from a finite set A_i. Second, each player i observes a private signal y_i from a finite set Y_i.[4] Third, each player i simultaneously sends a public message $m_i \in M_i$ to the other players. We assume that the message space M_i is countable and includes Y_i^l for each $l \geq 1$ so that player i can report any finite sequence of private signals in any period t.[5] Fourth, each player i simultaneously makes a publicly observable side-payment τ_{ij} to each player j.[6] Finally, the players observe χ, the outcome of a public randomization device, which is uniformly distributed between 0 and 1.[7]

Let $a = (a_1, \dots, a_n)$ denote an action profile and let $y = (y_1, \dots, y_n)$ denote a signal profile. Let a_{-i} and y_{-i} denote a minus a_i and y minus y_i, respectively.[8] Denote the set of action profiles by A and the set of private-signal profiles by Y. Conditional on a, each y is realized with probability $p(y|a)$. The marginal probabilities of y_{-i}, y_i, and (y_i, y_j) are denoted, respectively, by $p_{-i}(y_{-i}|a)$, $p_i(y_i|a)$, and $p_{ij}(y_i, y_j|a)$, and the marginal probabilities of y_{-i} and y_j, conditional on a and y_i, are denoted, respectively, by $p_{-i}(y_{-i}|a, y_i)$ and $p_j(y_j|a, y_i)$. We assume that the signal distribution has full support; that is, $p(y|a) > 0$ for each $y \in Y$ and $a \in A$. This assumption, among other things, rules out public monitoring.

[4] Each player i does not observe the others' actions. Instead, the signal y_i serves as an imperfect indicator of the joint actions taken.

[5] For any set X, X^l denotes the $l$ary Cartesian power of X.

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

[7] Since we assume that players can communicate, the public randomization device can be replaced by a series of jointly controlled lotteries and, hence, is dispensable. See Aumann et al. (1968).

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

Player i's gross payoff in the stage game is

$$r_i(a_i, y_i) + \sum_{j=1}^{n}(\tau_{ji} - \tau_{ij}).$$

We refer to $r_i(a_i, y_i)$ as player i's stage-game payoff. For each player i, the stage-game payoff function, $r_i(a_i, y_i)$, depends solely on a_i and y_i. The actions of the other players affect $r_i(a_i, y_i)$ only through the distribution of y_i. Player i's expected stage-game payoff conditional on a is

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

We say that an action profile a^* is efficient if it maximizes $\sum_{i=1}^{n} g_i(a)$. Let G_0 denote the stage game G without the last three steps (i.e., reporting private signals, making side-payments, and observing the outcome of the public randomization device). To save notation, we assume that G_0 has a pure-strategy Nash equilibrium and denote it by $a^N = (a_1^N, a_2^N, \dots, a_n^N)$.[9]

The players discount future payoffs by a common discount factor $\delta < 1$. Player i's average repeated-game payoff is

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

where $a_{i,t}$, $y_{i,t}$, and $\tau_{ij,t}$ are the period-t values of a_i, y_i, and τ_{ij}, respectively. At the beginning of each period t, each player i has observed a private history that consists of his actions and signals in the previous $(t-1)$ periods, as well as a public history that consists of the signal reports, side-payments, and outcomes of the public randomization device in the previous $(t-1)$ periods. We use $h_{i,t}$ to denote the history, both private and public, that player i observes at the beginning of period t. Let H_i denote the set of all finite histories for player i. A pure strategy $\sigma_i = (\alpha_i, \rho_i, b_i)$ for player i consists of three components: an action strategy α_i that maps each history $h_{i,t} \in H_i$ into an action in A_i, a reporting strategy ρ_i that maps each $(h_{i,t}, a_{i,t}, y_{i,t})$ into a message in M_i, and a transfer strategy $b_i = (b_{i1}, b_{i2}, \dots, b_{in})$ that maps each $(h_{i,t}, a_{i,t}, y_{i,t}, m_t)$ into an n-vector of nonnegative real numbers. Mixed strategies are defined in the standard way. We will state clearly if mixing is involved, but for simplicity, we do not introduce separate notations for mixed strategies. Following Compte (1998) and Obara (2009), we use the solution concept of perfect T-public equilibrium. A strategy of player i is T-public (for some $T \geq 1$) if the following conditions are satisfied for any positive integer l: (i) at any period t, where $lT < t \leq (l+1)T$, the stage-game action strategy depends only on the public history at lT and always prescribes the same action; (ii) at any period t, where $lT < t < (l+1)T$, the stage-game reporting strategy is always fully mixed and depends only on the public history at lT; (iii) at period $t = (l+1)T$, the stage-game reporting strategy depends only on player i's private signals after period lT. A strategy profile $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ is a

[9] All results go through if we replace a^N with a mixed-strategy Nash equilibrium of G_0.

perfect T-public equilibrium if each σ_i is T-public and the continuation strategy profile after any public history at periods $lT+1$, for any $l \geq 0$, constitutes a Nash equilibrium. Perfect T-public equilibrium is weaker than sequential equilibrium in that it does not require consistency and sequential rationality within a T block. However, given the full-support assumption, for any perfect T-public equilibrium, there exists a sequential equilibrium with the same equilibrium outcomes (Lemma 2, Kandori and Matsushima 1998).

3. Result

Approximate efficiency is attainable only if there is a way for the players to monitor each other. In the literature, the standard monitoring assumption is statistical detectability. Formally, unilateral deviations from an action profile a are statistically detectable if for each player i,

$$p_{-i}(\cdot|a) \notin \text{co}\big(\{p_{-i}(\cdot|a_i', a_{-i})|a_i' \neq a_i\}\big).$$

While the condition is generic mathematically when the number of signals is sufficiently large, it does not follow naturally from standard economic assumptions. For example, in the secret-price-cutting problem where each firm's private signal is its own demand, it is unclear what economic restrictions on the demand system would imply statistical detectability.

In this paper, we propose an alternative approach to achieve approximate efficiency. We say that an efficient action profile a^* is *locally unique* if for each player $i \in \mathcal{N}$ and each $a_i' \in A_i/\{a_i^*\}$,

$$\sum_{j=1}^{n} g_j(a_i', a_{-i}^*) < \sum_{j=1}^{n} g_j(a^*).$$

Our main result is the following theorem.

Theorem 1. *Suppose there is a locally unique efficient action profile a^*. Then for any $\epsilon > 0$, there exists $\bar{\delta} < 1$ such that for each $\delta > \bar{\delta}$, there is a perfect T-public equilibrium with total average equilibrium payoff greater than $\sum_{i=1}^{n} g_i(a^*) - \epsilon$.*

See Appendix B for the proof.

In many economic problems, it is natural to assume that players can observe their own payoffs and exchange side-payments. Theorem 1 shows that in such problems, statistical detectability is not necessary for approximate efficiency. If σ is an equilibrium, then it is also an equilibrium for the players to use a public randomization device to randomize between σ and playing a^N forever in the beginning of the game. Hence, given the conditions in Theorem 1, we could construct equilibrium with total average payoff anywhere strictly between $\sum_{i=1}^{n} g_i(a^*)$ and $\sum_{i=1}^{n} g_i(a^N)$. Furthermore, since we can use side-payments to redistribute payoffs, the set of average equilibrium payoff profiles as δ goes to 1 includes any payoff profile with total payoff strictly between $\sum_{i=1}^{n} g_i(a^N)$ and

$\sum_{i=1}^{n} g_i(a^*)$. In equilibrium, a player i with $g_i(a^*) < g_i(a^N)$ would receive payments from the other players to raise his average equilibrium payoff above $g_i(a^N)$.

When the players observe their own payoffs, statistical detectability is not necessary because we can enforce an efficient outcome by making the continuation payoff of each player equal to the total payoff of the other players.[10] To achieve approximate efficiency, we introduce a new solution to the learning problem in a delayed-communication mechanism with correlated signals. Given a_{-i}, there is a statistical test that can detect both deviations from a_i and false reports of y_i if

$$p_{-i}(\cdot|a, y_i) \notin \text{co}\big(\{p_{-i}(\cdot|a_i', a_{-i}, y_i')|(a_i', y_i') \neq (a_i, y_i)\}\big)$$

for all $y_i \in Y_i$. Obara (2009) shows that the learning problem can be solved if deviations and false reports are jointly statistically detectable for at least one player.[11] Our solution, instead, requires the existence of a locally unique efficient action profile. The requirement rules out unilateral deviations that change the payoff distribution for some player without changing the expected payoff of any player.

We prove Theorem 1 by construction. The availability of side-payments allows us to construct an equilibrium that involves only a^* and a^N. The equilibrium consists of two states: a cooperative state and a noncooperative state. The players start off in the cooperative state. In the cooperative state, the players play a^* for T periods and send completely mixed random messages in the first $(T-1)$ periods. At the end of period T, they reveal their private signals in all T periods and exchange side-payments. They then use a public randomization device to determine the state in the next period. If the players stay in the cooperative state, they repeat the same process and play a^* for another T periods. If they switch to the noncooperative state, they play the stage-game Nash equilibrium a^N and send fully mixed random messages forever.

4. Main issues

4.1 *A T-period game*

To analyze the players' incentives in the cooperative state, we consider the following T-period game, which we denote by $G^{T,\delta}(S)$. In each period $k = 1, \dots, T$, the players play the stage game G_0.[12] At the end of period T, the players simultaneously report the private signals they have received during the T periods. Let $\hat{y}_i^T = (\hat{y}_i(1), \dots, \hat{y}_i(T))$ denote a signal report of player i and let $\hat{y}^T = (\hat{y}_1^T, \dots, \hat{y}_n^T)$ denote a signal-report profile. In addition to his stage-game payoffs, each player i receives a transfer $S_i(\hat{y}^T)$ that depends on the signal-report profile $\hat{y}^T$. Player i's total discounted payoff in this T-period game is

$$\sum_{k=1}^{T} \delta^{k-1} r_i(a_i(k), y_i(k)) + S_i(\hat{y}^T),$$

where $a_i(k)$ is player i's period-k action and $y_i(k)$ is his period-k signal.

[10] The assumption that payoffs are observable is crucial. See Rahman (2012) for complications that would arise when this assumption fails.

[11] Obara (2009) calls such a player an informed player.

[12] Recall that G_0 denotes the stage game G without the last three steps.

Because there is no external source of payoffs in the original repeated game, we require that the total transfer be nonpositive; i.e.,

$$\sum_{i=1}^{n} S_i(\hat{y}^T) \le 0, \quad \text{for all } \hat{y}^T \in Y^T. \tag{1}$$

For any $k \le T$, let $y_i^k \equiv (y_i(1), \dots, y_i(k))$ denote player i's signals in the first k periods. Player i's strategy consists of two components: an action strategy α_i^T that maps each $y_i^k \in \bigcup_{l=0}^{T-1} Y_i^l$ into an action $a_i \in A_i$ and a reporting strategy ρ_i^T that maps each $y_i^T \in Y_i^T$ into a report $\hat{y}_i^T \in Y_i^T$.[13] Let $\mathcal{A}_i^T$ and Σ_i^T be player i's action–strategy set and reporting–strategy set, respectively. Denote the action strategy that chooses a_i^* in every period by α_i^{T*}, the truth-telling reporting strategy by ρ_i^{T*}, and the strategy profile where every player i chooses $(\alpha_i^{T*}, \rho_i^{T*})$ by (α^{T*}, ρ^{T*}). We say S enforces a^* if (α^{T*}, ρ^{T*}) is a Nash equilibrium. The total per-period efficiency loss of S conditional on (α^{T*}, ρ^{T*}) is

$$\mathrm{WL}(T, \delta, S) \equiv -\sum_{i=1}^{n} \frac{1-\delta}{1-\delta^T} E_{y^T}[S_i(y^T) | \alpha^{T*}].$$

To prove Theorem 1, we must show that for any $\epsilon > 0$, it is possible to find S that satisfies (1) and enforces a^* with efficiency loss less than ϵ.

4.2 *Reducing efficiency loss through delayed communication*

In this section, we review the independent-signal case. The result was first proved by Compte (1998). Our argument follows that of Fong et al. (2011).

Let

$$\Pi_j(y_j^T) \equiv \sum_{k=1}^{T} \delta^{k-1} r_j(a_j^*, y_j(k))$$

denote player j's discounted stage-game payoff in the T-period game. Hoeffding's inequality provides an upper bound on the probability that Π_j exceeds a certain value.

Lemma 1 (Hoeffding's inequality (Hoeffding 1963, Theorem 2)). *Let $x_1, x_2, \dots, x_{l_0}$ be independent random variables such that $|x_l| \le \nu$ for each $l \le l_0$. Then, for any $d > 0$, we have*

$$\Pr\left(\sum_{l=1}^{l_0} x_l \ge E\left[\sum_{l=1}^{l_0} x_l\right] + d\right) \le \exp\left(-\frac{d^2}{2\nu^2 l_0}\right). \tag{2}$$

Let $c_1 \equiv \max_{j, y_j} |r_j(a_j^*, y_j)|$. Substituting $\delta^{k-1} r_j(a_j^*, y_j(k))$ for x_l, $T^{2/3}$ for d, T for l_0, and c_1 for ν into (2), we have

$$\Pr\left(\left\{y_j^T : \Pi_j(y_j^T) > \sum_{k=1}^{T} \delta^{k-1} g_j(a^*) + T^{2/3}\right\} \Big| \alpha^{T*}\right) \le \exp\left(-\frac{1}{2c_1^2} T^{1/3}\right). \tag{3}$$

[13] As usual, y^0 denotes the null history $\varnothing$ and Y^0 denotes the set whose only element is y^0.

Although $T^{2/3}/\sum_{k=1}^{T}\delta^{k-1}g_j(a^*)$ tends to 0 as T goes to infinity and δ goes to 1, the probability that Π_j exceeds the mean by more than $T^{2/3}$ decreases exponentially in $T^{1/3}$. Intuitively, as T increases, the support of Π_j widens, but the distribution becomes increasingly concentrated around the mean.

Consider a vector of transfer functions $S^* \equiv (S_1^*, \dots, S_n^*)$. For each player i and each $\hat{y}^T \in Y^T$,

$$S_i^*(\hat{y}^T) \equiv -\sum_{j\neq i} \max\{K_j - \Pi_j(\hat{y}_j^T), 0\}, \tag{4}$$

where

$$K_j \equiv \sum_{k=1}^{T} \delta^{k-1} g_j(a^*) + T^{2/3}. \tag{5}$$

Under S^*, each player i's transfer is increasing one-to-one in the reported discounted payoff of each player $j \neq i$ up to a cap K_j, which is set at $T^{2/3}$ above the mean discounted payoff of player j when α^{T*} is chosen. It follows from (3) that $\mathrm{WL}(T, \delta, S^*)$ could be made arbitrarily small by making T sufficiently large and δ sufficiently close to 1.

The strategy profile (α^{T*}, ρ^{T*}) is a Nash equilibrium of $G^{T,\delta}(S^*)$ when the signals are independent. As S_i^* does not depend on ρ_i^T, the truthful reporting strategy ρ_i^{T*} is a best response. For each $y_i^k \in \bigcup_{l=0}^{T} Y_i^l$, let

$$U_i^T(\alpha_i^T, \rho_i^T; S, y_i^k) \equiv E_{y^T}\left[\sum_{l=1}^{T} \delta^{l-1} r_i(\alpha_i^T(y_i^{l-1}), y_i(l)) + S_i(\rho_i^T(y_i^T), y_{-i}^T)\Big|\alpha_i^T, \alpha_{-i}^{T*}, y_i^k\right] \tag{6}$$

denote player i's expected payoff conditional on y_i^k when he chooses (α_i^T, ρ_i^T) and other players choose $(\alpha_{-i}^{T*}, \rho_{-i}^{T*})$. Write $U_i^T(\alpha_i^T, \rho_i^T; S)$ for $U_i^T(\alpha_i^T, \rho_i^T; S, y_i^0)$. Substitute (4) and (5) into (6) (with S^* replacing S) and rearrange terms. We have

$$U_i^T(\alpha_i^T, \rho_i^T; S^*, y_i^k) = V_i(\alpha_i^T; y_i^k) - R_i(\alpha_i^T; y_i^k) - \sum_{j\neq i} K_j,$$

where

$$V_i(\alpha_i^T; y_i^k) \equiv E_{y^T}\left[\sum_{l=1}^{T} \delta^{l-1}\left(r_i(\alpha_i^T(y^{l-1}), y_i(l)) + \sum_{j\neq i} r_j(a_j^*, y_j(l))\right)\Big|\alpha_i^T, \alpha_{-i}^{T*}, y_i^k\right]$$

$$R_i(\alpha_i^T; y_i^k) \equiv \sum_{j\neq i} E_{y^T}\left[\max\{\Pi_j(y_j^T) - K_j, 0\}|\alpha_i^T, \alpha_{-i}^{T*}, y_i^k\right].$$

The variable V_i is the players' total expected payoff conditional on y_i^k when player i chooses α_i^T and other players choose α_{-i}^{T*}. Because a^* is locally unique, any deviation from α_i^{T*} must strictly lower V_i. Let Δ denote the minimum loss in total expected stage-game payoff when a player deviates unilaterally from a^*, and let $\mathcal{A}_i^T(y_i^k)$ denote the set of player i's action strategies that prescribe $a_i \neq a_i^*$ after y_i^k in period $(k+1)$. For any $k < T$, $y_i^k \in Y_i^k$, and $\alpha_i^T \in \mathcal{A}_i^T(y_i^k)$,

$$V_i(\alpha_i^{T*}; y_i^k) - V_i(\alpha_i^T; y_i^k) \geq \delta^k \Delta.$$

The variable R_i is the part of the payoffs of players $j \neq i$ not captured by S_i^*. It measures the effect of truncating player i's incentives. When the signals are independent, player i believes that the truncation effect is very small throughout the mechanism. By (3),

$$\Pr\left(\left\{y_j^T : \Pi_j(y_j^T) > \sum_{k=1}^{T} \delta^{k-1} g_j(a^*) + T^{2/3}\right\} \Big| \alpha^{T*}, y_i^k\right) \leq \exp\left(-\frac{1}{2c_1^2} T^{1/3}\right).$$

Since $\Pi_j(y_j^T) \leq c_1 T$ and $R_i(\alpha_i^T; y_i^k) \geq 0$,

$$R_i(\alpha_i^{T*}; y_i^k) - R_i(\alpha_i^T; y_i^k) \leq c_1 T \exp\left(-\frac{1}{2c_1^2} T^{1/3}\right),$$

which would be less than $\delta^k \Delta$ when T is large and δ is close to 1. Hence, α_i^{T*} is a best response against $(\alpha_{-i}^{T*}, \rho_{-i}^{T*})$ when T is sufficiently large and δ is sufficiently close to 1.

4.3 *Solving the learning problem*

When the signals are correlated, we can no longer make each player i believe that the truncation does not matter throughout the T-period game. Even when ex ante it is extremely unlikely that player j's discounted payoff will be greater than K_j, there is always some small chance that player i may learn from his own signals during the T-period game that the probability is not negligible, and when that happens, the fact that S_i^* depends on Π_j only up to K_j would affect player i's incentive to choose a_i^* in the remaining periods.

The main contribution of this paper is to provide a solution to this learning problem. The basic idea is to make player j pay player i a reward that cancels out the truncation effect whenever player i thinks there is a nonnegligible probability that Π_j is greater than K_j. However, we cannot directly make player j pay player i whenever Π_j is larger than K_j, because doing so would merely transfer player i's learning problem to player j—player j may now want to deviate from a_j^* when Π_j is large. To avoid this problem, we make the additional reward dependent on a proxy variable that is uncorrelated with player j's private signals but very likely to be large whenever Π_j is large from the perspective of player i.

We first introduce the proxy variable. For any i, $j \neq i$ and any $(y_i, y_j) \in Y_i \times Y_j$, let

$$z_{ij}(y_i, y_j) \equiv E_{y_j'}[r_j(a_j^*, y_j')|a^*, y_i] \frac{p_i(y_i|a^*) p_j(y_j|a^*)}{p_{ij}(y_i, y_j|a^*)}.$$

Note that z_{ij} is always defined due to the full-support assumption. By Bayes' rule,

$$\frac{p_i(y_i|a) p_i(y_j|a)}{p_{ij}(y_i, y_j|a)} = \frac{p_i(y_i|a)}{p_i(y_i|y_j, a)} = \frac{p_j(y_j|a)}{p_j(y_j|y_i, a)}.$$

Hence, for any $(y_i, y_j) \in Y_i \times Y_j$,

$$E_{y_j}[z_{ij}(y_i, y_j)|a^*, y_i] = E_{y_j'}[r_j(a_j^*, y_j')|a^*, y_i] \tag{7}$$

and

$$E_{y_i}[z_{ij}(y_j, y_i)|a^*, y_j] = E_{y'_j}[r_j(a^*_j, y'_j)|a^*]. \tag{8}$$

These properties extend to the total discounted value of z_{ij} over the T-period game. For any $(y_i^T, y_j^T) \in Y_i^T \times Y_j^T$, let

$$\Gamma_{ij}(y_i^T, y_j^T) \equiv \sum_{k=1}^{T} \delta^{k-1} z_{ij}(y_i(k), y_j(k)).$$

It follows immediately from (7) and (8) that for any $k \leq T$ and $(y_i^k, y_j^k) \in Y_i^k \times Y_j^k$,

$$E_{y_i^T, y_j^T}[\Gamma_{ij}(y_i^T, y_j^T)|\alpha^{T*}, y_i^k] = E_{y_i^T, y_j^T}[\Pi_j(y_j^T)|\alpha^{T*}, y_i^k] \tag{9}$$

and

$$E_{y_i^T, y_j^T}[\Gamma_{ij}(y_i^T, y_j^T)|\alpha^{T*}, y_j^k] = E_{y_i^T, y_j^T}[\Pi_j(y_j^T)|\alpha^{T*}]. \tag{10}$$

Equations (9) and (10) mean that throughout the T-period game, player i's conditional expectation of Γ_{ij} is always equal to his conditional expectation of Π_j, while player j's conditional expectation of Γ_{ij} is always equal to the unconditional expectation of Π_j.

We add two extra components to S^* to deal with the learning problem. The first extra component for player i is

$$\begin{aligned} L_i(\hat{y}^T) \equiv & \sum_{j \neq i} \max\{\Pi_j(\hat{y}_j^T) - K_j, 0\} f_{ij}(\hat{y}^T) \\ & - \sum_{j \neq i} \max\big\{(\bar{\Pi}_i - K_i) - \max\{K_j - \Pi_j(\hat{y}_j^T), 0\}, 0\big\} f_{ji}(\hat{y}^T), \end{aligned} \tag{11}$$

where $\bar{\Pi}_j \equiv \max_{\hat{y}_j^T} \Pi_j(\hat{y}_j^T)$ and

$$f_{ij}(\hat{y}^T) \equiv \begin{cases} 1 & \text{if } \Gamma_{ij}(\hat{y}_i^T, \hat{y}_j^T) > K_j - \frac{1}{2}T^{2/3} \\ 0 & \text{otherwise.} \end{cases}$$

Intuitively, for any player pair ij, we raise player i's transfer by $\max\{\Pi_j(\hat{y}_j^T) - K_j, 0\}$ when $\Gamma_{ij}(\hat{y}_i^T, \hat{y}_j^T)$ is greater than $(K_j - \frac{1}{2}T^{2/3})$ and simultaneously reduce player j's transfer by

$$\max\big\{(\bar{\Pi}_j - K_j) - \max\{K_i - \Pi_i(\hat{y}_i^T), 0\}, 0\big\}.$$

The first summation term on the right-hand side of (11) is the total extra reward player i receives, and the second summation term is the total reward he pays. Note that while we say player j "pays" player i a reward, player j actually pays more than what player i receives.

The amount deducted from player j's payoff is made independent of $\hat{y}_j^T$ to ensure that player j has no incentive to misreport to reduce the amount he needs to pay. (See Remark 1 below for more details.) Nevertheless, player i may still try to manipulate L_i

by lying about his signals, as f_{ij} and f_{ji} both depend on $\hat{y}_i^T$. To induce player i to report truthfully, we add a second component,

$$D_i(\hat{y}^T) \equiv T^{-2} \sum_{k=1}^{T} \log p_{-i}(\hat{y}_{-i}(k) | a^*, \hat{y}_i(k)),$$

to player i's transfer. The variable D_i is always negative and it is bounded from below as the signal distribution has full support.

With these two extra components, player i's total transfer is now

$$S_i^{**}(\hat{y}^T) = S_i^*(\hat{y}^T) + L_i(\hat{y}^T) + D_i(\hat{y}^T). \tag{12}$$

Let $S^{**} = (S_1^{**}, \dots, S_n^{**})$. Since for any $\hat{y}^T \in Y^T$,

$$\sum_{i=1}^{n} (S_i^*(\hat{y}^T) + L_i(\hat{y}^T)) \leq 0 \quad \text{and} \quad \sum_{i=1}^{n} D_i(\hat{y}^T) \leq 0,$$

the total transfer is always negative.

Lemma 2. *For any $\epsilon > 0$, there exists a T_0 such that, for all $T \geq T_0$ and all $\delta \geq 1 - T^{-2}$, the following statements hold:*

(i) We have $\text{WL}(T, \delta, S^{**}) \leq \epsilon$.

(ii) The strategy profile (α^{T*}, ρ^{T*}) is a Nash equilibrium of $G^{T,\delta}(S^{**})$.

See Appendix A for the proof.

Part (i) of Lemma 2 comes from the fact that the expected values of both extra components are very small when T is large and (α^{T*}, ρ^{T*}) is chosen. As T goes to infinity, D_i goes to zero. The expected value of $L_i(y^T)$ is small because for any player j, ex ante it is unlikely for Γ_{ji} to be greater than $(K_i - \frac{1}{2}T^{2/3})$ when T is large and (α^{T*}, ρ^{T*}) is chosen.

We turn to part (ii) of Lemma 2. Substituting (12) into (6) and rearranging terms, we have

$$U_i^T(\alpha_i^T, \rho_i^T; S^{**}, y_i^k) = V_i(\alpha_i^T; y_i^k) - \sum_{l=1}^{3} R_i^l(\alpha_i^T, \rho_i^T; y_i^k) - \sum_{j \neq i} K_j, \tag{13}$$

where

$$R_i^1(\alpha_i^T, \rho_i^T; y_i^k) \equiv \sum_{j \neq i} E_{y^T}\big[\max\{\Pi_j(y_j^T) - K_j, 0\}\big(1 - f_{ij}(\rho_i^T(y_i^T), y_{-i}^T)\big) | \alpha_i^T, \alpha_{-i}^{T*}, y_i^k\big]$$

$$R_i^2(\alpha_i^T, \rho_i^T; y_i^k) \equiv \sum_{j \neq i} E_{y^T}\big[(\bar{\Pi}_i - K_i) f_{ji}(\rho_i^T(y_i^T), y_{-i}^T) | \alpha_i^T, \alpha_{-i}^{T*}, y_i^k\big]$$

$$R_i^3(\alpha_i^T, \rho_i^T; y_i^k) \equiv -E_{y^T}\big[D_i(\rho_i^T(y_i^T), y_{-i}^T) | \alpha_i^T, \alpha_{-i}^{T*}, y_i^k\big].$$

Note that R_i^1, R_i^2, and R_i^3 are always positive.

The term R_i^1 in (13) is the net truncation effect. Recall that S_i^* increases linearly in player j's payoff up to K_j. Since the first component of L_i pays player i an additional reward $\max\{\Pi_j(\hat{y}_j^T) - K_j, 0\}$ when $f_{ij}(\hat{y}^T) = 1$, the cap K_j binds for player i only when $\Pi_j > K_j$ and $f_{ij}(\hat{y}^T) = 0$. Under truthful reporting, the event occurs when

$$\Gamma_{ij}(y_i^T, y_j^T) + \tfrac{1}{2}T^{2/3} < K_j < \Pi_j(y_j^T). \tag{14}$$

By (9), for any $k \leq T$ and $y_i^k \in Y_i^k$,

$$E_{y_i^T, y_j^T}[\Gamma_{ij}(y_i^T, y_j^T) - \Pi_j(y_j^T) | \alpha^{T*}, y_i^k] = 0.$$

As T goes to infinity,

$$\Pr\Big(\Big\{(y_i^T, y_j^T) : \Gamma_{ij}(y_i^T, y_j^T) - \Pi_j(y_j^T) < -\tfrac{1}{2}T^{2/3}\Big\} | \alpha^{T*}, y_i^k\Big)$$

converges to 0. Hence, $R_i^1(\alpha_i^{T*}, \rho_i^{T*}; y_i^k)$ is extremely small when T is large. Roughly speaking, player i believes he will almost always receive the additional reward when $\Pi_j > K_j$ if he chooses $(\alpha_i^{T*}, \rho_i^{T*})$ and other players choose $(\alpha_{-i}^{T*}, \rho_{-i}^{T*})$.

The second component of L_i deducts

$$\max\big\{(\bar{\Pi}_i - K_i) - \max\{K_j - \Pi_j(\hat{y}_j^T), 0\}, 0\big\}$$

from player i's payoff whenever Γ_{ji} is greater than $(K_i - \frac{1}{2}T^{2/3})$ for some $j \neq i$. The term R_i^2 in (13) denotes the expected amount deducted. By (10),

$$E_{y_j^T, y_i^T}[\Gamma_{ji}(y_j^T, y_i^T) | \alpha^{T*}, y_i^k] = E_{y_i^T}[\Pi_i(y_i^T) | \alpha^{T*}] = K_i - T^{2/3}$$

for any $k \leq T$ and $y_i^k \in Y_i^k$. As T becomes large, the probability that Γ_{ji} is greater than $(K_i - \frac{1}{2}T^{2/3})$ is extremely low. Therefore, $R_i^2(\alpha_i^{T*}, \rho_i^{T*}; y_i^k)$ is extremely small when T is large. In other words, player i believes he will almost never have to pay the reward to player j when (α^{T*}, ρ^{T*}) is chosen.

We have argued that $R_i^1(\alpha_i^{T*}, \rho_i^{T*}; y_i^k)$ and $R_i^2(\alpha_i^{T*}, \rho_i^{T*}; y_i^k)$ are both extremely small when T is large. Because D_i is bounded from below, $R_i^3(\alpha_i^{T*}, \rho_i^{T*}; y_i^k)$ converges to zero as T goes to infinity. Since R_i^1, R_i^2, and R_i^3 are always positive, as T becomes large, for any strategy (α_i^T, ρ_i^T) and any $y_i^k \in Y_i^k$,

$$\sum_{l=1}^{3}(R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) - R_i^l(\alpha_i^T, \rho_i^T; y_i^k))$$

cannot be significantly greater than 0. Hence, using an argument similar to the independent-signal case, we can show that when T is sufficiently large and δ sufficiently close to 1,

$$V_i(\alpha_i^{T*}; y_i^k) - V_i(\alpha_i^T; y_i^k) \geq \sum_{l=1}^{3} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) - \sum_{l=1}^{3} R_i^l(\alpha_i^T, \rho_i^T; y_i^k)$$

for any $y_i^k \in Y_i^k$, $\alpha_i^T \in \mathcal{A}_i^T(y_i^k)$, and $\rho_i^T \in \Sigma_i^T$. It is, therefore, a best response for player i to choose α_i^{T*} when other players choose $(\alpha_{-i}^{T*}, \rho_{-i}^{T*})$.

Suppose player i chooses α_i^{T*}. Player i can lower R_i^1 and R_i^2 only by lying about his posterior beliefs in some period l; i.e., reporting $\hat{y}_i(l)$ with $p_{-i}(\cdot|a^*, \hat{y}_i(l)) \neq p_{-i}(\cdot|a^*, y_i(l))$. The function D_i is a scoring rule. It is well known that for any $y_i^T \in Y_i^T$, $\hat{y}_i^T$ maximizes

$$E_{y_{-i}^T}[D_i(\tilde{y}_i^T, y_{-i}^T)|\alpha^{T*}, y_i^T]$$

with respect to $\tilde{y}_i^T$ if and only if $p_{-i}(\cdot|a^*, \hat{y}_i(l)) = p_{-i}(\cdot|a^*, y_i(l))$ for any $l \leq T$. If player i lies about his posterior beliefs, R_i^3 will strictly increase. Since under truth-telling, R_i^1 and R_i^2 converge to zero exponentially faster than R_i^3, any decrease in R_i^1 and R_i^2 will be dominated by the increase in R_i^3 when T is sufficiently large.

Remark 1. We raise player i's payoff by $\max\{\Pi_j(\hat{y}_j^T) - K_j, 0\}$ and reduce player j's payoff by

$$\max\big\{(\bar{\Pi}_j - K_j) - \max\{K_i - \Pi_i(\hat{y}_i^T), 0\}, 0\big\} \tag{15}$$

when $f_{ij}(\hat{y}^T) = 1$. Since $(\bar{\Pi}_j - K_j) \geq \max\{\Pi_j(\hat{y}_j^T) - K_j, 0\}$, the total payoff of the players may be reduced as a result. This does not affect the approximate-efficiency result, because it is extremely unlikely that $f_{ij}(\hat{y}^T) = 1$ in equilibrium.

Note that we cannot simply reduce player j's payoff by

$$\max\big\{\max\{\Pi_j(\hat{y}_j^T) - K_j, 0\} - \max\{K_i - \Pi_i(\hat{y}_i^T), 0\}, 0\big\} \tag{16}$$

because doing so may give player j incentives to lie about y_j^T. The component D_j ensures only that player j would be strictly worse off if he lies about his posterior beliefs. However, there may exist y_j and y_j' such that $p_{-j}(\cdot|a^*, y_j) = p_{-j}(\cdot|a^*, y_j')$ and $r_j(a_j^*, y_j) > r_j(a_j^*, y_j')$. In that case, player j would strictly prefer to report y_j' when he observes y_j if his payoff were reduced by (16) when $f_{ij}(\hat{y}^T) = 1$. More generally, the amount deducted from player j when $f_{ij}(\hat{y}^T) = 1$ must satisfy three conditions: (i) the amount must be big enough so that the total transfer remains nonpositive, (ii) the amount cannot depend on player j's report, and (iii) the amount deducted must always be positive. It is straightforward to see that (15) satisfies all three conditions. To see why the last requirement is necessary, note that when the amount deducted from player j when $f_{ij}(\hat{y}^T) = 1$ is positive, player j's incentive is to reduce the chance that $f_{ij}(\hat{y}^T) = 1$. Since the probability $f_{ij}(\hat{y}^T) = 1$ is already extremely close to zero in equilibrium, the potential gain from any deviation is extremely small. If the amount deducted were negative (meaning that player j will be rewarded when $f_{ij}(\hat{y}^T) = 1$), player j's incentive would be to increase the chance that $f_{ij}(\hat{y}^T) = 1$.

REMARK 2. Let $\lambda = (\lambda_1, \dots, \lambda_n)$ denote an n-vector. Consider the constrained maximization problem

$$Q(\lambda, T, \delta) \equiv \max_{\alpha^T, \rho^T, S} \frac{1-\delta}{1-\delta^T} \sum_{i=1}^{n} \lambda_i U_i^T(\alpha_i^T, \rho_i^T; S)$$

$$\text{subject to} \sum_{i=1}^{n} \lambda_i S_i(y^T) \leq 0 \quad \text{for all } y^T \in Y^T$$

$$(\alpha^T, \rho^T) \text{ is a Nash equilibrium of } G^{T,\delta}(S).$$

Lemma 2 says that for any λ such that $\lambda_1 = \cdots = \lambda_n$, the value of $Q(\lambda, T, \delta)$ is ϵ-close to $\sum_{i=1}^{n} g_i(a^*)$ when T is large and δ is close to 1. The same technique can be used to provide an ϵ-tight bound on the value of $Q(\lambda, T, \delta)$ for any $\lambda \gg 0$ when T is large and δ is close to 1.[14] The technique, however, does not apply for λ with $\lambda_i < 0$ for some i. Recall that we argue that since R_i^1, R_i^2, and R_i^3 are close to 0 in equilibrium, no deviation could reduce these terms significantly. For the argument to work, R_i^1, R_i^2, and R_i^3 must be always positive. If $\lambda_i < 0$ for some i, then we would need to reward player i by giving him a positive transfer whenever the payoffs of the other players are above certain thresholds. In that case, the "truncation" terms would be negative.

5. WITHOUT SIDE-PAYMENTS

Following the influential work of Fudenberg and Levine (1994), it is common to prove folk-theorem-type results by finding a tight lower bound on the value of $Q(\lambda, T, \delta)$ for all λ with $|\lambda| = 1$. Since our technique does not apply to λ where $\lambda_i < 0$ for some i, we cannot apply their result.

Nevertheless, our method of solving the learning problem may still apply when side-payments are not allowed. We say that a pure action profile a is locally unique if for each player i, there is a vector $\lambda^{a,i} = (\lambda_1^{a,i}, \lambda_2^{a,i}, \dots, \lambda_n^{a,i})$ with $\lambda_i^{a,i} > 0$ such that

$$\sum_{j=1}^{n} \lambda_j^{a,i} g_j(a) > \sum_{j=1}^{n} \lambda_j^{a,i} g_j(a_i', a_{-i}) \quad \text{for each } a_i' \in A_i. \tag{17}$$

Let α^{Ta} denote the action strategy profile in the T-period game where the players always choose a in each period. Using the techniques developed in the last section, we can show that for any locally unique action profile a and any $\epsilon > 0$, there exists a T_0 such that, for all $T \geq T_0$ and all $\delta \geq 1 - T^{-2}$, there is a transfer-function profile $\widetilde{S}^a = (\widetilde{S}_1^a, \dots, \widetilde{S}_n^a)$ that enforces a in the T-period game with $\sum_i \widetilde{S}_i^a(y^T) \leq 0$ for all y^T and

$$\left| \frac{1-\delta}{1-\delta^T} E_{y^T}[\widetilde{S}_i^a(y^T) | \alpha^{Ta}] \right| < \epsilon$$

for each player i.

[14] See Appendix C for details.

Let $\mathcal{B}$ denote a set of locally unique action profiles that contains a^*. Consider a T-public trigger-strategy profile that contains $\#\mathcal{B}$ cooperative states, each corresponding to some $a \in \mathcal{B}$. The strategy profile starts off in cooperative state a^*. In each cooperative state $a \in \mathcal{B}$, the players play a for T periods, send completely mixed random messages during the first $(T-1)$ periods, and reveal their private signals in all T periods at the end of period T. Conditional on the report profile $\hat{y}^T$, the players switch to each state $a' \in \mathcal{B}$ in the next period with probability $\mu^a_{a'}(\hat{y}^T)$ and the noncooperative state with probability $\mu^a_N(\hat{y}^T)$. Once in the noncooperative state, the players play a^N and send completely mixed random messages forever.

Let $v^{Ta} = (v^{Ta}_1, \dots, v^{Ta}_n)$ denote the average discounted payoff at the beginning of state $a \in \mathcal{B}$, and let $v^{TN} = (v^N_1, \dots, v^N_n)$ denote the average discounted payoff at the beginning of the noncooperative state. By construction, $v^{TN} = g(a^N)$ for all T. The strategy profile is a perfect T-public equilibrium if the transition probabilities $\{\mu^a_{a'}\}_{a' \in \mathcal{B} \cup \{N\}}$ are chosen such that for all $a \in \mathcal{B}$ and $y^T \in Y^T$,

$$\frac{\delta^T}{1-\delta} \sum_{a' \in \mathcal{B} \cup \{N\}} \mu^a_{a'}(y^T)(v^{Ta'} - v^{Ta^*}) = \widetilde{S}^a(y^T). \tag{18}$$

Such transition probabilities exist if and only if, for any $a \in \mathcal{B}$,

$$\left\{ \frac{1-\delta}{1-\delta^T} \delta^{-T} \widetilde{S}^a(y^T) \middle| y^T \in Y^T \right\} \subseteq \mathrm{co}\left(\left\{ \frac{v^{Ta'} - v^{Ta^*}}{1-\delta^T} \middle| a' \in \mathcal{B} \cup \{N\} \right\} \right). \tag{19}$$

The left-hand side of (19) is the set of per-period transfers needed to enforce a in a T-period block, while the right-hand side is the set of feasible per-period transfers that can be carried out by switching to a different state at the end of the block.

Both the left-hand and right-hand sides of (19) vary with T and δ. Suppose (18) holds for each $a \in \mathcal{B}$ and $y^T \in Y^T$. Then

$$v^{Ta}_i = (1-\delta^T)(g_i(a) + \kappa^a_i(T, \delta)) + \delta^T v^{Ta^*}_i,$$

with

$$\kappa^a_i(T, \delta) \equiv \frac{(1-\delta) E_{y^T}[\widetilde{S}^a_i(y^T) | \alpha^{Ta}]}{1-\delta^T}.$$

As δ goes to 1, we can choose $T(\delta)$ such that $\kappa^a_i(T(\delta), \delta)$ goes to 0 and $\delta^{T(\delta)}$ goes to 1. The right-hand side of (19) would then be approximately equal to

$$W(\mathcal{B}) \equiv \bigcup_{\xi \in \mathbb{R}_{++}} \mathrm{co}\big(\{w(a)\}_{a \in \mathcal{B}} \cup \{\xi w(a^N)\}\big),$$

where

$$w(a) \equiv g(a) - g(a^*).$$

When $\kappa^a_i(T(\delta), \delta)$ is close to 0 and $\delta^{T(\delta)}$ is close to 1, the efficiency loss in enforcing a becomes negligible, and switching from cooperative state a^* to cooperative state a is equivalent to switching from a^* to a for T periods. Since a switch to a^N is permanent, $w(a^N)$ carries an arbitrarily large weight in $W(\mathcal{B})$.

For any $i, j \in \mathcal{N}$, $i \neq j$, let

$$W^{ij}(\mathcal{B}) \equiv \{(x_i, x_j) \in \mathbb{R}^2 \mid \exists x' \in W(\mathcal{B}) : x'_i = x_i, x'_j = x_j, x'_k = 0 \text{ for } k \neq i, j\}$$

denote the subset of $W(\mathcal{B})$ with zero net transfer to all players other than i and j. Given the full-support assumption, we may need to reward player i and punish player j without changing the continuation payoffs of the other players. Hence, the trigger-strategy profile could be an equilibrium only if, for any distinct player pair ij, $W^{ij}(\mathcal{B})$ contains a point with a positive i component and a point with a positive j component.

The following theorem provides sufficient conditions for the trigger-strategy profile to be an approximately efficient equilibrium.

Theorem 2. *For each locally unique Pareto-efficient action profile a and each distinct player pair ij, let*

$$\overline{d}_{ij}^a \equiv \frac{n(n-1)}{2}(\lambda_i^{a,i})^{-1}\lambda_j^{a,i}\Big(\max_{y_j} r_j(a_j, y_j) - g_j(a)\Big) \tag{20}$$

$$\underline{d}_{ij}^a \equiv \frac{n(n-1)}{2}(\lambda_i^{a,i})^{-1}\lambda_j^{a,i}\Big(\min_{y_j} r_j(a_j, y_j) - g_j(a)\Big). \tag{21}$$

Suppose there is a set of locally unique Pareto-efficient action profiles $\mathcal{B}$ that contains the efficient profile a^* such that for any players $i, j \in \mathcal{N}$ with $i \neq j$, and any $a \in \mathcal{B}$, there exist $\overline{w}_{ij}^a$, $\overline{w}_{ji}^a$, $\underline{w}_{ij}^a$, and $\underline{w}_{ij}^a$ such that

$$\overline{w}_{ij}^a > \overline{d}_{ij}^a, \qquad \underline{w}_{ij}^a < \underline{d}_{ij}^a, \qquad \overline{w}_{ji}^a > \overline{d}_{ji}^a, \qquad \underline{w}_{ji}^a < \underline{d}_{ji}^a \tag{22}$$

and

$$(\overline{w}_{ij}^a, \underline{w}_{ji}^a), (\underline{w}_{ij}^a, \overline{w}_{ji}^a), (\underline{w}_{ij}^a, \underline{w}_{ji}^a) \in \operatorname{int}(W^{ij}(\mathcal{B})). \tag{23}$$

Then for any $\epsilon > 0$, there exists $\bar{\delta} < 1$ such that for each $\delta > \bar{\delta}$, there is a perfect T-public equilibrium without side-payments in which the average equilibrium payoff of each player i is greater than $g_i(a^*) - \epsilon$.

See Appendix C for the proof.

Note that transferring payoffs through actions is generally inefficient; i.e., it costs player i more than 1 unit of payoff to raise the payoff of player j by 1 unit. Nevertheless, the equilibrium is approximately efficient because the total expected transfer goes to 0 as δ goes to 1.

Under our transfer scheme, player i receives a reward when player j's payoff is above average and pays a penalty when player j's payoff is below average. Roughly speaking,

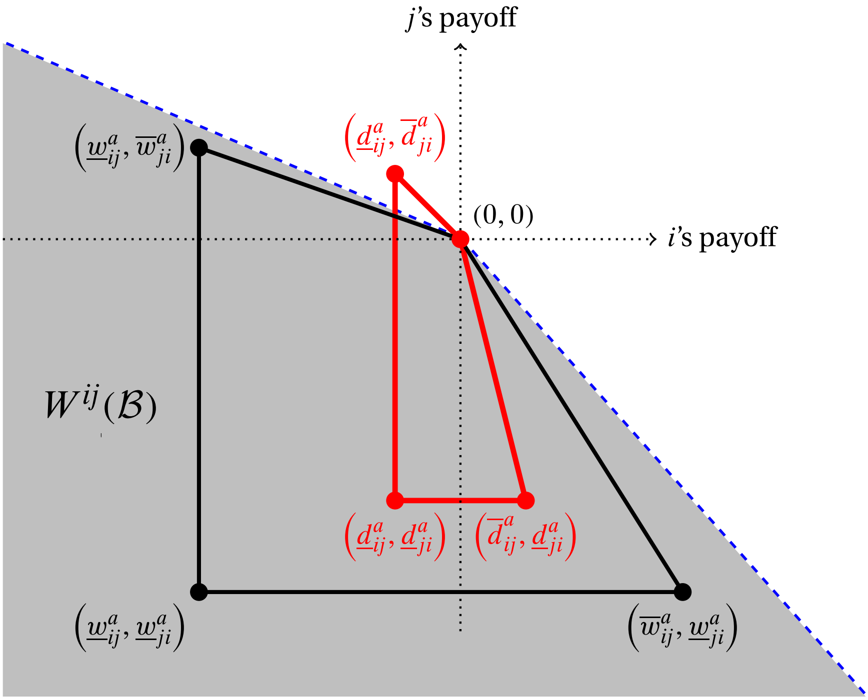

FIGURE 1. Illustration of conditions (22) and (23). The shaded area denotes $W^{ij}(\mathcal{B})$.

$\overline{d}^a_{ij}$ and $\underline{d}^a_{ij}$ are, respectively, the maximum and minimum of the part of player i's transfer tied to j's payoffs when a is enforced, multiplied by $\frac{1}{2}n(n-1)$.[15] Conditions (22) and (23) require that $W^{ij}(\mathcal{B})$ be large enough so that, first, when player i (j) is rewarded by $\overline{d}^a_{ij}$ ($\overline{d}^a_{ji}$), it is possible to punish player j (i) by $\underline{d}^a_{ji}$ ($\underline{d}^a_{ij}$) or more and, second, it is possible punish player i by $\underline{d}^a_{ij}$ and player j by $\underline{d}^a_{ji}$ simultaneously. See Figure 1. The condition ensures the transfers between players i and j can be implemented with a "probability quota" of $2/(n(n-1))$. Since there are $\frac{1}{2}n(n-1)$ distinct pairs of players, we can implement the transfers between every pair of players simultaneously. Note that the left-hand side of (23) depends on the impact of a player's private signal on his own payoff given the action profile, while the right-hand side depends on the impact of different action profiles on the expected payoffs of the players. Our approach is, therefore, more likely to apply without side-payments when actions have a greater impact on the payoffs than do private signals.

To illustrate the restrictions imposed by conditions (22) and (23), consider the following noisy prisoners' dilemma game. The action sets are $A_1 = A_2 = \{0, 1\}$ and the signal sets are $Y_1 = Y_2 \subset \mathbb{R}$. Normalize $\min\{y_i \in Y_i\} = 0$ and denote $\max\{y_i \in Y_i\}$ by $\overline{y}$. The stage-game payoff for player $i = 1, 2$ is

$$r_i(a_i, y_i) = a_i + y_i.$$

[15]These terms are multiplied by $(\lambda^{a,i}_i)^{-1}\lambda^{a,i}_j$ to reflect the different weights assigned to the payoffs of players i and j in (17).

The expectation of y_i depends only on a_{-i} and not a_i.[16] For each $i = 1, 2$,

$$E[y_i|a_{-i} = 0] = \phi_0\overline{y}$$
$$E[y_i|a_{-i} = 1] = \phi_1\overline{y},$$

where $\phi_0, \phi_1 > 0$ and $(\phi_0 - \phi_1)\overline{y} > 1$. The expected stage-game payoffs are

$$g(0,0) = (\phi_0\overline{y}, \phi_0\overline{y})$$
$$g(1,1) = (1 + \phi_1\overline{y}, 1 + \phi_1\overline{y})$$
$$g(0,1) = (\phi_1\overline{y}, 1 + \phi_0\overline{y})$$
$$g(1,0) = (1 + \phi_0\overline{y}, \phi_1\overline{y}).$$

The unique stage-game Nash equilibrium is $(1, 1)$ and the efficient outcome is $(0, 0)$.

Let $\mathcal{B} \equiv \{(0,0), (0,1), (1,0)\}$. For two-player games, $W^{12}(\mathcal{B}) = W^{21}(\mathcal{B}) = W(\mathcal{B})$. It is straightforward to check that $(x_1, x_2) \in \operatorname{int}(W(\mathcal{B}))$ if and only if

$$(\phi_0 - \phi_1)\overline{y}x_1 + x_2 < 0$$
$$x_1 + (\phi_0 - \phi_1)\overline{y}x_2 < 0$$
$$x_1 - x_2 < 1 + (\phi_0 - \phi_1)\overline{y}$$
$$-x_1 + x_2 < 1 + (\phi_0 - \phi_1)\overline{y}.$$

See Figure 2.

Since $a_1 = 0$ strictly maximizes $g_1(a_1, 1) + g_2(a_1, 1)$ and $a_2 = 1$ strictly maximizes $g_2(0, a_2)$, $(a_1, a_2) = (0, 1)$ is locally unique with $\lambda^{(0,1),1} = (1, 1)$ and $\lambda^{(0,1),2} = (0, 1)$. Similarly, $(a_1, a_2) = (1, 0)$ is locally unique with $\lambda^{(1,0),1} = (1, 0)$ and $\lambda^{(1,0),2} = (1, 1)$.

Note that for $a_2 \in \{0, 1\}$ and $a_1 = 0$,

$$\max_{y_2} r_2(a_2, y_2) - g_2(0, a_2) = (1 - \phi_0)\overline{y}$$
$$\min_{y_2} r_2(a_2, y_2) - g_2(0, a_2) = -\phi_0\overline{y}.$$

Hence, in any state $a \in \mathcal{B}$, player 1's maximum reward and punishment are $(1 - \phi_0)\overline{y}$ and $-\phi_0\overline{y}$, respectively. Since the game is symmetric, the same holds for player 2. Following (20) and (21), we have

a	$\overline{d}^a_{12}$	$\underline{d}^a_{12}$	$\overline{d}^a_{21}$	$\underline{d}^a_{21}$
0, 0	$(1 - \phi_0)\overline{y}$	$-\phi_0\overline{y}$	$(1 - \phi_0)\overline{y}$	$-\phi_0\overline{y}$
0, 1	$(1 - \phi_0)\overline{y}$	$-\phi_0\overline{y}$	0	0
1, 0	0	0	$(1 - \phi_0)\overline{y}$	$-\phi_0\overline{y}$

[16]Note that this does not imply that y_1 and y_2 are uncorrelated conditional on (a_1, a_2).

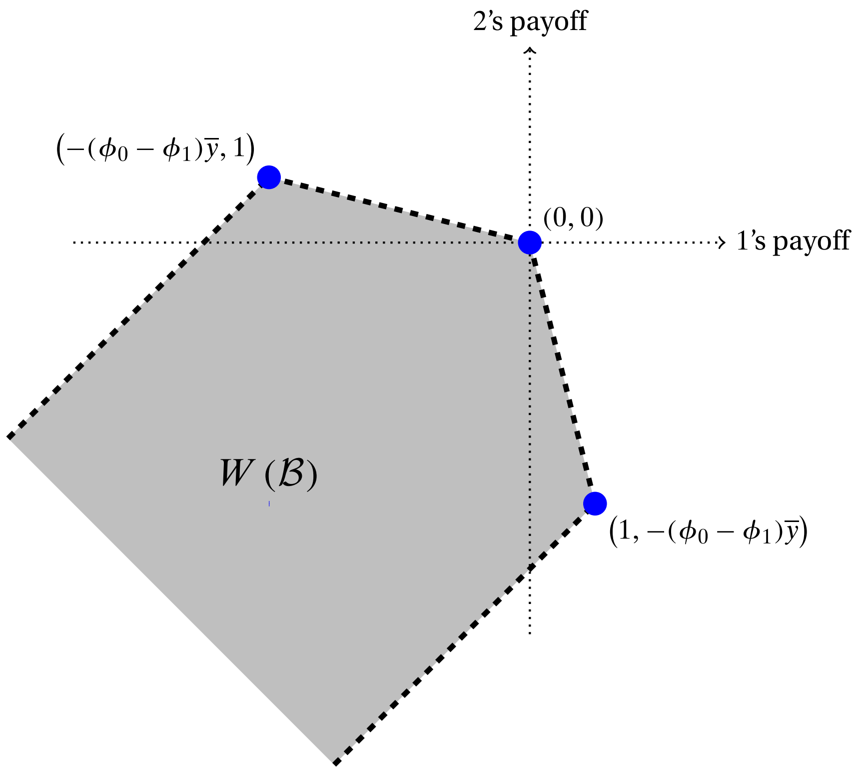

FIGURE 2. The set $W(\mathcal{B})$.

For any $\varepsilon > 0$, setting

$$\overline{w}_{ij}^a = \overline{w}_{ji}^a > (1-\phi_0)\overline{y} + \varepsilon$$
$$\underline{w}_{ij}^a = \underline{w}_{ji}^a < -\phi_0\overline{y} - \varepsilon$$

would satisfy condition (22) for any $a \in \mathcal{B}$. Conditions (22) and (23), therefore, can be satisfied for each $a \in \mathcal{B}$ if

$$((1-\phi_0)\overline{y}, -\phi_0\overline{y}), (-\phi_0\overline{y}, (1-\phi_0)\overline{y}), (-\phi_0\overline{y}, -\phi_0\overline{y}) \in \operatorname{int}(W(\mathcal{B})).$$

It is clear from Figure 2 that $(-\phi_0\overline{y}, -\phi_0\overline{y}) \in \operatorname{int}(W(\mathcal{B}))$. For $((1-\phi_0)\overline{y}, -\phi_0\overline{y})$ and $(-\phi_0\overline{y}, (1-\phi_0)\overline{y})$ to belong to $\operatorname{int}(W(\mathcal{B}))$, we need

$$(1-\phi_0)\overline{y} < 1 \tag{24}$$
$$\overline{y} < (\phi_0 - \phi_1)\overline{y} + 1. \tag{25}$$

Condition (24) ensures that the maximum reward is less than the maximum amount that can be transferred to player 1 by switching from $(0,0)$ to $(0,1)$. Condition (25) ensures that it is possible to implement the punishment for player 2 while paying the maximum reward to player 1. Since $(\phi_0 - \phi_1)\overline{y} > 1$, these conditions are satisfied when $\overline{y}$, the width of the support of the private signals, is less than $\min\{2, 1/(1-\phi_0)\}$.

6. Conclusion

We introduce a new solution to the learning problem associated with delayed communication. Using this solution, we show that in economic problems where players can observe their own payoffs, statistical detectability is not necessary for approximate efficiency.

We establish approximate efficiency with side-payments by constructing a trigger-strategy equilibrium in which only a^* and a^N are played. Since our technique requires negative transfers, we cannot extend our result to the no-side-payment case by applying the results of Fudenberg and Levine (1994). In Section 5, we show that approximate efficiency may still hold without side-payments if the players can carry out the necessary transfers by switching between locally unique action profiles. However, for this approach to work, the transfers required must be relatively small.

A key step of our analysis is to show that for each player pair ij, it is possible to construct a proxy variable that is highly correlated with player i's expectation of player j's payoff but uncorrelated with player j's signal. The construction depends on the assumption that the signal distribution has full support. The assumption ensures that no player can use his own signal to rule out any signal of another player. It is well known that approximate efficiency may be unattainable when the signals are public (Radner et al. 1986 and Abreu et al. 1991). In future work we plan to investigate whether our method can be extended to intermediate cases between private monitoring with full support and imperfect public monitoring.

Appendix A: Proof of Lemma 2

We first derive a bound for each $R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k)$. Let

$$c_2 = \max_{i,j,y_i,y_j} |r_j(a_j^*, y_j) - z_{ij}(y_i, y_j)|.$$

Note that from (9), we have

$$E_{y_i^T, y_j^T}[\Pi_j(y_j^T) - \Gamma_{ij}(y_i^T, y_j^T) | \alpha^{T*}, y_i^k] = 0.$$

Hence, substituting $\delta^{l-1}(r_j(a_j^*, y_j(l)) - z_{ij}(y_i(l), y_j(l)))$ for x_l, $\frac{1}{2}T^{2/3}$ for d, T for l_0, and c_2 for ν in Hoeffding's inequality, we have

$$\Pr\left(\left\{(y_i^T, y_j^T) : \Pi_j(y_j^T) - \Gamma_{ij}(y_i^T, y_j^T) > \frac{1}{2}T^{2/3}\right\} \middle| \alpha^{T*}, y_i^k\right) \le \exp\left(-\frac{1}{8c_2^2}T^{1/3}\right).$$

Since (14) holds only when $\Pi_j - \Gamma_{ij} > \frac{1}{2}T^{2/3}$ and since $\max\{\Pi_j(\hat{y}_j^T) - K_j, 0\} \le c_1 T$,

$$R_i^1(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) \le (n-1)c_1 T \exp\left(-\frac{1}{8c_2^2}T^{1/3}\right). \tag{26}$$

Let $c_3 = \max_{i,j,y_i,y_j} |z_{ij}(y_i, y_j)|$. Since the expected value of Γ_{ji} conditional on y_i^k is equal to $(K_i - T^{2/3})$, substituting $\delta^{l-1} z_{ji}(y_i(l), y_j(l))$ for x_l, $T^{2/3}$ for d, T for l_0, and c_3 for ν in

Hoeffding's inequality, we have

$$\Pr\big(\{(y_i^T, y_j^T) : \Gamma_{ji}(y_i^T, y_j^T) > K_i\} | \alpha^{T*}, y_i^k\big) \le \exp\left(-\frac{1}{2c_3^2} T^{1/3}\right).$$

It follows that

$$R_i^2(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) \le (n-1)c_1 T \exp\left(-\frac{1}{2c_3^2} T^{1/3}\right). \tag{27}$$

Let $c_4 = \max_{y,i} |\log p_{-i}(y_{-i}|a^*, y_i)|$. By definition we have

$$D_i(y^T) = T^{-2} \sum_{k=1}^{T} \log p_{-i}(y_{-i}(k)|a^*, y_i(k)) \ge -c_4 T^{-1}.$$

Hence,

$$R_i^3(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) \le c_4 T^{-1}. \tag{28}$$

Combining all these bounds and noting that $R_i^l(\alpha_i^T, \rho_i^T; y_i^k) \ge 0$ for each l, we have

$$\begin{aligned} &\sum_{l=1}^{3} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) - \sum_{l=1}^{3} R_i^l(\alpha_i^T, \rho_i^T; y_i^k) \\ &\quad \le (n-1)c_1 T \exp\left(-\frac{1}{8c_2^2} T^{1/3}\right) + (n-1)c_1 T \exp\left(-\frac{1}{2c_3^2} T^{1/3}\right) + c_4 T^{-1}. \end{aligned} \tag{29}$$

We have argued in the main text that, for any $\alpha_i^T \in \mathcal{A}^T(y_i^k)$,

$$V_i(\alpha_i^{T*}; y_i^k) - V_i(\alpha_i^T; y_i^k) \ge \delta^k \Delta.$$

Set $\delta^*(T) \equiv 1 - T^{-2}$. This ensures that $(\delta^*(T))^T$ tends to 1 as T tends to infinity. Note that the right-hand side of (29) tends to 0 as T tends to infinity. So we can choose T_1 large enough such that for all $T \ge T_1$ and $\delta \ge \delta^*(T)$,

$$\begin{aligned} \delta^T \Delta &\ge (\delta^*(T))^T \Delta \\ &\ge (n-1)c_1 T \exp\left(-\frac{1}{8c_2^2} T^{1/3}\right) + (n-1)c_1 T \exp\left(-\frac{1}{2c_3^2} T^{1/3}\right) + c_4 T^{-1}. \end{aligned}$$

It follows that

$$V_i(\alpha_i^{T*}; y_i^k) - V_i(\alpha_i^T; y_i^k) \ge \sum_{l=1}^{3} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) - \sum_{l=1}^{3} R_i^l(\alpha_i^T, \rho_i^T; y_i^k).$$

Thus, for $(\alpha_i^T, \rho_i^T) \in \mathcal{A}^T(y_i^k) \times \Sigma_i^T$, $y_i^k \in Y_i^k$ and $k = 0, 1, \ldots, (T-1)$,

$$U_i^T(\alpha_i^{T*}, \rho_i^{T*}; S^{**}, y_i^k) \ge U_i^T(\alpha_i^T, \rho_i^T; S^{**}, y_i^k). \tag{30}$$

To prove that $(\alpha_i^{T*}, \rho_i^{T*})$ maximizes $U_i^T(\alpha_i^T, \rho_i^T; S^{**})$ with respect to (α_i^T, ρ_i^T), we still need to show that, for all $\rho_i^T \in \Sigma_i^T$,

$$U_i^T(\alpha_i^{T*}, \rho_i^{T*}; S^{**}, y_i^T) \geq U_i^T(\alpha_i^{T*}, \rho_i^T; S^{**}, y_i^T). \tag{31}$$

When player i follows α_i^{T*} the reporting strategy ρ_i^T affects only R_i^l, $l = 1, 2, 3$, but not V_i. Write $\hat{y}_i^T$ for $\rho_i^T(y_i^T)$. There are two cases to consider.

Case 1. Suppose $p_{-i}(\cdot|a^*, \hat{y}_i(k)) = p_{-i}(\cdot|a^*, y_i(k))$ for each $k \in \{1, \ldots, T\}$. In this case, for $l = 1, 2, 3$,

$$R_i^l(\alpha_i^{T*}, \rho_i^T; y_i^T) = R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^T).$$

The case for $l = 3$ follows immediately from the definition of D_i. To see that the values of R_i^1 and R_i^2 are also the same, note that

$$E_{y_j'(k)}[r_j(a_j^*, y_j'(k))|a^*, y_i(k)] = E_{y_j'(k)}[r_j(a_j^*, y_j'(k))|a^*, \hat{y}_i(k)]$$

and

$$\begin{aligned} \frac{p_i(y_i(k)|a^*)p_j(y_j(k)|a^*)}{p_{ij}(y_i(k), y_j(k)|a^*)} &= \frac{p_j(y_j(k)|a^*)}{p_j(y_j(k)|a^*, y_i(k))} = \frac{p_j(y_j(k)|a^*)}{p_j(y_j(k)|a^*, \hat{y}_i(k))} \\ &= \frac{p_i(\hat{y}_i(k)|a^*)p_j(y_j(k)|a^*)}{p_{ij}(\hat{y}_i(k), y_j(k)|a^*)}. \end{aligned}$$

It follows that $z_{ij}(y_i(k), y_j(k)) = z_{ij}(\hat{y}_i(k), y_j(k))$ and $z_{ji}(y_i(k), y_j(k)) = z_{ji}(\hat{y}_i(k), y_j(k))$ for each $y_j(k)$. Since R_i^1 and R_i^2 depend only on z_{ij}, z_{ji}, and other players' reports, they have the same value under y_i^T and under $\hat{y}_i^T$.

Case 2. Suppose $p_{-i}(\cdot|a^*, \hat{y}_i(k)) \neq p_{-i}(\cdot|a^*, y_i(k))$ for some $k \in \{1, \ldots, T\}$. It is a standard result in the scoring-rule literature that

$$\begin{aligned} &E_{y_{-i}(k)}[\log(p_{-i}(y_{-i}(k)|a^*, \hat{y}_i(k)))|a^*, y_i(k)] \\ &\qquad < E_{y_{-i}(k)}[\log(p_{-i}(y_{-i}(k)|a^*, y_i(k)))|a^*, y_i(k)]. \end{aligned}$$

Hence,

$$R_i^3(\alpha_i^{T*}, \rho_i^T; y_i^T) - R_i^3(\alpha_i^{T*}, \rho_i^{T*}; y_i^T) \geq T^{-2}d,$$

where

$$\begin{aligned} d \equiv \min\bigl\{&E_{y_{-i}}[\log(p_{-i}(y_{-i}|a^*, y_i))|a^*, y_i] - E_{y_{-i}}[\log(p_{-i}(y_{-i}|a^*, \hat{y}_i))|a^*, y_i] \mid \\ &y_i, \hat{y}_i \in Y_i \text{ and } p_{-i}(\cdot|a^*, \hat{y}_i) \neq p_{-i}(\cdot|a^*, y_i)\bigr\} > 0. \end{aligned}$$

We have already shown that

$$\begin{aligned} &\sum_{l=1}^{2} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) - \sum_{l=1}^{2} R_i^l(\alpha_i^{T*}, \rho_i^T; y_i^k) \\ &\qquad \leq (n-1)c_1 T \exp\left(-\frac{1}{8c_2^2}T^{1/3}\right) + (n-1)c_1 T \exp\left(-\frac{1}{2c_3^2}T^{1/3}\right). \end{aligned} \tag{32}$$

Since the right-hand side of (32) decays faster than $T^{-2}d$, we can choose T_2 large enough such that for all $T \geq T_2$ and all δ,

$$R_i^3(\alpha_i^{T*}, \rho_i^T; y_i^T) - R_i^3(\alpha_i^{T*}, \rho_i^{T*}; y_i^T) \geq \sum_{l=1}^{2} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^k) - \sum_{l=1}^{2} R_i^l(\alpha_i^{T*}, \rho_i^T; y_i^k).$$

We now turn to the first part of the lemma. By definition we have

$$\begin{aligned} E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] &= \sum_{j\neq i} E_{y^T}[\Pi_j(y_j^T) - K_j|\alpha^{T*}] - \sum_{l=1}^{3} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^0) \\ &= -(n-1)T^{2/3} - \sum_{l=1}^{3} R_i^l(\alpha_i^{T*}, \rho_i^{T*}; y_i^0). \end{aligned}$$

Since $(\delta^*(T))^T$ tends to 1 as T tends to infinity, $(1-\delta^*(T))(1-(\delta^*(T))^T)^{-1}$ is of the order of T^{-1}. It follows that

$$\frac{1-\delta^*(T)}{1-(\delta^*(T))^T}(n-1)T^{2/3}$$

tends to 0 as T tends to infinity. By (26), (27), and (28), each R_i^l is bounded by a term that tends to 0 as T tends to infinity. Hence, for any ϵ, we can choose T_3 large enough such that for all $T \geq T_3$ and $\delta \geq \delta^*(T)$, and for each i,

$$-\frac{1-\delta}{1-\delta^T} E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] < \frac{\epsilon}{n}. \tag{33}$$

Set $T_0 = \max\{T_1, T_2, T_3\}$. Then (30) and (31) hold for all $T \geq T_0$ and $\delta \geq \delta^*(T)$, and by (33), $\text{WL}(T, \delta, S^{**}) \leq \epsilon$.

Appendix B: Proof of Theorem 1

Theorem 1 is obviously true when $\sum_{i=1}^n g_i(a^*) = \sum_{i=1}^n g_i(a^N)$. Henceforth, we assume that $\sum_{i=1}^n g_i(a^*) > \sum_{i=1}^n g_i(a^N)$. We prove Theorem 1 by constructing a Nash-threat trigger-strategy equilibrium with the desired properties. The basic structure of the equilibrium strategy profile has already been given in Section 3.

To complete the description of the equilibrium strategy profile, we need to specify the side-payments and the transition probabilities after T periods in the cooperative state. Fix $\epsilon > 0$. By Lemma 2, we can pick T_0 so that when $T \geq T_0$ and $\delta \geq \delta^*(T)$,

$$\text{WL}(T, \delta, S^{**}) < \min\left\{\epsilon, \sum_{i=1}^n g_i(a^*) - \sum_{i=1}^n g_i(a^N)\right\}. \tag{34}$$

Pick a vector $\zeta = (\zeta_1, \zeta_2, \ldots, \zeta_n)$ with $\sum_{i=1}^n \zeta_i = 0$ such that for each player i,

$$g_i(a^*) + \frac{1-\delta}{1-\delta^T} E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] + \zeta_i > g_i(a^N).$$

Pick $\bar{\delta} \in (\delta^*(T), 1)$ such that for each player i, each $y^T \in Y^T$ and each $\delta \geq \bar{\delta}$,

$$S_i^{**}(y^T) + \frac{1-\delta^T}{1-\delta}\zeta_i \geq -\frac{\delta^T}{1-\delta}\left(g_i(a^*) + \frac{1-\delta}{1-\delta^T}E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] + \zeta_i - g_i(a^N)\right). \tag{35}$$

Let $\mu(y^T)$ denote the probability of switching to the noncooperative state after the players report y^T at the end of period T in the cooperative state, and let $\beta_{ij}(y^T) \geq 0$ denote the amount player i pays player j. We define $\mu(y^T)$ and $\beta_{ij}(y^T)$ as follows. First, for any $y^T \in Y^T$, set

$$\mu(y^T) \equiv \frac{-\delta^{-T}(1-\delta)\sum_{i=1}^n S_i^{**}(y^T)}{\sum_{i=1}^n (g_i(a^*) + \frac{1-\delta}{1-\delta^T}E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] - g_i(a^N))}. \tag{36}$$

By (34) and (35), $\mu(y^T) \in [0, 1]$ for all y^T.

Next, set

$$\beta_{ij}(y^T) \equiv \begin{cases} \beta_i^{\text{net}}(y^T)\frac{\min(\beta_j^{\text{net}}(y^T),0)}{\sum_{k=1}^n \min(\beta_k^{\text{net}}(y^T),0)} & \text{if } \beta_i^{\text{net}}(y^T) > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\begin{aligned}\beta_i^{\text{net}}(y^T) \equiv & \left(-S_i^{**}(y^T) - \frac{1-\delta^T}{1-\delta}\zeta_i\right)\delta^{-T+1} \\ & -(1-\delta)^{-1}\delta\mu(y^T)\left(g_i(a^*) + \frac{1-\delta}{1-\delta^T}E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] + \zeta_i - g_i(a^N)\right).\end{aligned} \tag{37}$$

When $\beta_i^{\text{net}}(y^T) > 0$, $\sum_{j=1}^n \beta_{ij}(y^T) = \beta_i^{\text{net}}(y^T)$ and $\sum_{j=1}^n \beta_{ji}(y^T) = 0$. When $\beta_i^{\text{net}}(y^T) \leq 0$, $\sum_{j=1}^n \beta_{ij}(y^T) = 0$ and $\sum_{j=1}^n \beta_{ji}(y^T) = -\beta_i^{\text{net}}(y^T)$; the last follows from the fact that, by (36) and (37), $\sum_{i=1}^n \beta_i^{\text{net}}(y^T) = 0$ for all $y^T \in Y^T$. Hence, for any player i and any $y^T \in Y^T$, the net side-payment received by player i is always equal to

$$\sum_{j=1}^n (\beta_{ji}(y^T) - \beta_{ij}(y^T)) = -\beta_i^{\text{net}}(y^T). \tag{38}$$

Let v_i^* and v_i^N denote, respectively, player i's average discounted payoff at the beginning of the cooperative state and that at the beginning of the noncooperative state under this trigger-strategy profile. It is straightforward to see that $v_i^N = g_i(a^N)$. By standard arguments and (38), we have

$$\begin{aligned} v_i^* = & (1-\delta^T)g_i(a^*) - E_{y^T}[\beta_i^{\text{net}}(y^T)|\alpha^{T*}]\delta^{T-1}(1-\delta) \\ & + \left(E_{y^T}[\mu(y^T)|\alpha^{T*}](v_i^N - v_i^*) + v_i^*\right)\delta^T. \end{aligned} \tag{39}$$

Substituting (37) into (39) and rearranging terms, we have

$$v_i^* = g_i(a^*) + \frac{1-\delta}{1-\delta^T}E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] + \zeta_i. \tag{40}$$

We claim that the trigger-strategy profile constitutes a perfect T-public equilibrium when T is sufficiently large and δ is sufficiently close to 1. It is obvious that the continuation strategies in the noncooperative state are self-enforcing. Focus on the cooperative state. Given report profile y^T, the discounted continuation payoff of player i at the end of a T-period block in the cooperative state before side-payments are made is equal to

$$-\beta_i^{\text{net}}(y^T) + (\mu(y^T)(v_i^N - v_i^*) + v_i^*)\frac{\delta}{1-\delta},$$

which, by (37) and (40), is equivalent to

$$\left(S_i^{**}(y^T) + \frac{1-\delta^T}{1-\delta}\zeta_i\right)\delta^{-T+1} + v_i^*\frac{\delta}{1-\delta}.$$

According to the trigger-strategy profile, the players will switch to the noncooperative state with probability 1 if any player fails to make the required side-payments. It follows from (35) that when δ is sufficiently close to 1, it is a best response for player i to make the required side-payments.

Assume player i is to make the equilibrium side-payments at the end of period T. When players other than i follow the trigger-strategy profile, player i's discounted payoff for choosing a T-period action strategy α_i^T and a reporting strategy ρ_i^T, at the beginning of the cooperative state, is equal to

$$U_i^T(\alpha_i^T, \rho_i^T; S_i^{**}) + \frac{1-\delta^T}{1-\delta}\zeta_i + v_i^*\frac{\delta^T}{1-\delta}.$$

It follows from Lemma 2 that when T is sufficiently large and δ is sufficiently close to 1, it is a best response for player i to play the trigger-strategy profile within a T-period block in the cooperative state. This proves that the trigger-strategy profile is a perfect T-public equilibrium when T is sufficiently large and δ is sufficiently close to 1.

Finally, each player i's average continuation payoff at the beginning of the cooperative state, hence, at the beginning of the game, is

$$g_i(a^*) + \frac{1-\delta}{1-\delta^T}E_{y^T}[S_i^{**}(y^T)|\alpha^{T*}] + \zeta_i.$$

It follows from (34) that the total expected equilibrium payoff is greater than $\sum_{i=1}^n g_i(a^*) - \epsilon$.

Appendix C: Proof of Theorem 2

We prove the theorem in two steps.

Step 1. Enforcing the cooperative states in a T-period game. We first introduce transfer functions that implement $a \in \mathcal{B}$ in a T-period game with efficiency loss less than ϵ. Fix $a \in \mathcal{B}$ and $i, j \in \mathcal{N}$, $i \neq j$. Let

$$K_{ij}^a \equiv \sum_{k=1}^{T} \delta^{k-1}(\lambda_i^{a,i})^{-1}\lambda_j^{a,i} g_j(a) + T^{2/3}$$

$$\Pi^a_{ij}(y^T_j) \equiv \sum_{k=1}^{T} \delta^{k-1}(\lambda^{a,i}_i)^{-1}\lambda^{a,i}_j r_j(a_j, y_j(k))$$

$$z^a_{ij}(y_i, y_j) \equiv E_{y'_j}[(\lambda^{a,i}_i)^{-1}\lambda^{a,i}_j r_j(a_j, y'_j)|a, y_i]\frac{p_i(y_i|a)p_j(y_j|a)}{p_{ij}(y_i, y_j|a)}.$$

The second variable is defined for each $y^T_j \in Y^T_j$, and the third variable is defined for each $(y_i, y_j) \in Y_i \times Y_j$. These variables are the counterparts of K_j, Π_j, and z_{ij}, modified to reflect the different weights assigned to the players' stage-game payoff functions in (17). For any $\hat{y}^T \in Y^T$, let

$$\Gamma^a_{ij}(\hat{y}^T_i, \hat{y}^T_j) \equiv \sum_{k=1}^{T} \delta^{k-1} z^a_{ij}(\hat{y}_i(k), \hat{y}_j(k))$$

$$f^a_{ij}(\hat{y}^T) \equiv \begin{cases} 1 & \text{if } \Gamma^a_{ij}(\hat{y}^T_i, \hat{y}^T_j) > K^a_{ij} - \frac{1}{2}T^{2/3} \\ 0 & \text{otherwise.} \end{cases}$$

We can now introduce the transfer functions. For each $a \in \mathcal{B}$, each player i, and each $\hat{y}^T \in Y^T$, set the transfer for player i to be

$$\widetilde{S}^a_i(\hat{y}^T) \equiv S^a_i(\hat{y}^T) + \sum_{j\neq i} \widetilde{L}^a_{ij}(\hat{y}^T) + D^a_i(\hat{y}^T),$$

where

$$S^a_i(\hat{y}^T) \equiv -\sum_{j\neq i} \max\{K^a_{ij} - \Pi^a_{ij}(\hat{y}^T_j), 0\}$$

$$\begin{aligned} \widetilde{L}^a_{ij}(\hat{y}^T) \equiv{}& \max\{\Pi^a_{ij}(\hat{y}^T_j) - K^a_{ij}, 0\} f^a_{ij}(\hat{y}^T) \\ &- \max\left\{\left|\frac{\underline{w}^a_{ij}}{\overline{w}^a_{ji}}\right|\left(\max_{y^T_i} \Pi^a_{ji}(y^T_i) - K^a_{ji}\right) - \max\{K^a_{ij} - \Pi^a_{ij}(\hat{y}^T_j), 0\}, 0\right\} f^a_{ji}(\hat{y}^T) \end{aligned}$$

$$D^a_i(\hat{y}^T) \equiv T^{-2}\sum_{k=1}^{T} \log p_{-i}(\hat{y}_{-i}(k)|a, \hat{y}_i(k)).$$

Compared to the original transfer function S^{**}_i, the only difference is that in the definition of $\widetilde{L}^a_{ij}$, an extra term $|\underline{w}^a_{ij}/\overline{w}^a_{ji}|$ is inserted in the second maximum term on the right-hand side. This means that when $f^a_{ji}(\hat{y}^T) = 1$ (i.e., when player j receives an extra transfer), the payoff of player i is reduced by an amount equal to

$$\left|\frac{\underline{w}^a_{ij}}{\overline{w}^a_{ji}}\right|\left(\max_{y^T_i} \Pi^a_{ji}(y^T_i) - K^a_{ji}\right).$$

Since a^* maximizes total expected payoff, $|\underline{w}^a_{ij}/\overline{w}^a_{ji}| \geq 1$. When transferring payoffs through actions is inefficient, we need to lower the payoff for player i by more than 1 unit for every unit transferred to player j. Since this extra term is a constant, it does not affect the incentives of the players. Let α_i^{Ta} be player i's action strategy of always choosing a_i in the T-period game and let ρ^{T*} be the truthful reporting strategy. Following the argument behind Lemma 2, we have the following lemma.

Lemma 3. *For any $\epsilon > 0$, there exists a T_0 such that, for any $T \geq T_0$, $\delta \geq 1 - T^{-2}$, and $a \in \mathcal{B}$, the following statements hold:*

(i) We have $|((1-\delta)/(1-\delta^T))E_{y^T}[\widetilde{S}_i^a(y^T)|\alpha^{Ta}]| \leq \epsilon$ for each i.

(ii) The strategy profile (α^{Ta}, ρ^{T*}) is a Nash equilibrium of $G^{T,\delta}(\widetilde{S}^a)$.

Step 2. Implementing $\widetilde{S}^a$ in the repeated game. We described the equilibrium strategy profile in Section 5. Let

$$\kappa_i^a(T,\delta) \equiv \frac{1-\delta}{1-\delta^T} E_{y^T}[\widetilde{S}_i^a(y^T)|\alpha^{Ta}]$$

denote player i's expected per-period transfer in state a. Recall that $w(a) \equiv g(a) - g(a^*)$. Following the discussion in Section 5, to prove Theorem 2, we need to show that when δ is sufficiently large, we can choose T such that for all $a \in \mathcal{B}$ and $y^T \in Y^T$,

$$\begin{aligned} &\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}^a(y^T) \\ &\qquad \in W^{T,\delta}(\mathcal{B}) \qquad (41) \\ &\qquad \equiv \operatorname{co}\Big(\{w(a) + \kappa^a(T,\delta) - \kappa^{a^*}(T,\delta)\}_{a\in\mathcal{B}} \cup \Big\{\frac{w(a^N) - \kappa^{a^*}(T,\delta)}{1-\delta^T}\Big\}\Big). \end{aligned}$$

We first show that we can approximate $W^{T,\delta}(\mathcal{B})$ by $W(\mathcal{B})$ when T is sufficiently large and δ is sufficiently close to 1.

Lemma 4. *For any $x \in \operatorname{int}(W(\mathcal{B}))$, there exists T_1 such that for each $T \geq T_1$ and $\delta \geq 1 - T^{-2}$, $x \in W^{T,\delta}(\mathcal{B})$.*

Proof. Suppose the lemma were not true. Then there would exist a sequence $(T(1),\delta(1)),\dots,(T(l),\delta(l)),\dots$ such that for all $l \geq 1$, we have $T(l+1) > T(l)$, $\delta(l) \geq 1-(T(l))^{-2}$, and $x \notin W^{T(l),\delta(l)}(\mathcal{B})$. Since each $W^{T(l),\delta(l)}(\mathcal{B})$ is convex, by the theorem of separating hyperplanes, for each l there would exist $\theta(l) \in \mathbb{R}^n$ with $\sum_{k=1}^n(\theta_k(l))^2 = 1$ such that for all $a \in \mathcal{B}$,

$$\theta(l)\cdot x > \theta(l)\cdot\big(w(a) + \kappa^a(T(l),\delta(l)) - \kappa^{a^*}(T(l),\delta(l))\big) \qquad (42)$$

and

$$\theta(l)\cdot x > \theta(l)\cdot\bigg(\frac{w(a^N) - \kappa^{a^*}(T(l),\delta(l))}{1-(\delta(l))^{T(l)}}\bigg). \qquad (43)$$

Since each $\theta(l)$ is contained in the $(n-1)$-dimensional unit sphere, there is a subsequence converging to some θ in the unit sphere. Since, for each $i \in \mathcal{N}$ and $a \in \mathcal{B}$,

$\kappa_i^a(T(l),\delta(l))$ tends to 0 as l tends to infinity, from (42) we have for each $a\in\mathcal{B}$,

$$\theta\cdot x\geq\theta\cdot w(a).$$

Since $w(a^*)=0$, we have $\theta\cdot x\geq 0$. Since the left-hand side of (43) is bounded from above for any $\theta(l)$ and since $1-(\delta(l))^{T(l)}$ goes to 0 as l goes to infinity, (43) can hold for all l only if $\theta\cdot w(a^N)\leq 0$. Thus, we have for any $\xi\geq 0$,

$$\theta\cdot x\geq\theta\cdot(\xi w(a^N)).$$

This means that θ defines a hyperplane that separates x and $W(\mathcal{B})$, contradicting the assumption that x is an interior point of $W(\mathcal{B})$. □

Our next step is to show that we can implement the transfers between each pair of players i and j with a "probability quota" of $2/n(n-1)$ using only the points

$$(\overline{w}_{ij}^a,\underline{w}_{ji}^a),\qquad(\underline{w}_{ij}^a,\overline{w}_{ji}^a),\qquad(\underline{w}_{ij}^a,\underline{w}_{ji}^a),\qquad(0,0).$$

For any $i,j\in\mathcal{N}$, $i\neq j$, and for any $a\in\mathcal{B}$ and $y^T\in Y^T$, define

$$\widetilde{S}_{ij}^a(y^T)\equiv-\max\{K_{ij}^a-\Pi_{ij}^a(y_j^T),0\}+\widetilde{L}_{ij}^a(y^T)+\frac{1}{n-1}D_i^a(y^T).$$

By construction,

$$\widetilde{S}_i^a(y^T)=\sum_{j\neq i}\widetilde{S}_{ij}^a(y^T).$$

Lemma 5. *For any $a\in\mathcal{B}$, there exists T_2 such that, for any $T\geq T_2$, $\delta\geq 1-T^{-2}$, and $y^T\in Y^T$,*

$$\frac{n(n-1)}{2}\frac{1-\delta}{1-\delta^T}\delta^{-T}(\widetilde{S}_{ij}^a(y^T),\widetilde{S}_{ji}^a(y^T))\in\operatorname{co}(\{(\overline{w}_{ij}^a,\underline{w}_{ji}^a),(\underline{w}_{ij}^a,\overline{w}_{ji}^a),(\underline{w}_{ij}^a,\underline{w}_{ji}^a),(0,0)\}).\quad(44)$$

Proof. We prove the case for $a=a^*$. Other cases can be established in the same way. Notice that in this case $\Pi_{ij}^{a^*}(y_j^T)=\Pi_j(y_j^T)$ and $K_{ij}^{a^*}=K_j$. To simplify notation, we drop the superscript a^* in the variables (e.g., we write $\widetilde{S}_{ij}$ for $\widetilde{S}_{ij}^{a^*}$). It is clear from Figure 1 that (44) is equivalent to the conditions

$$\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}_{ij}(y^T)\geq\frac{2}{n(n-1)}\underline{w}_{ij}\quad(45)$$

$$\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}_{ji}(y^T)\geq\frac{2}{n(n-1)}\underline{w}_{ji}\quad(46)$$

$$\widetilde{S}_{ij}(y^T)-\frac{\overline{w}_{ij}}{\underline{w}_{ji}}\widetilde{S}_{ji}(y^T)\leq 0\quad(47)$$

$$-\frac{\overline{w}_{ji}}{\underline{w}_{ij}}\widetilde{S}_{ij}(y^T)+\widetilde{S}_{ji}(y^T)\leq 0.\quad(48)$$

Without loss of generality, let $i = 1$ and $j = 2$. Consider inequality (45). By the definition of $\widetilde{L}_{12}$,

$$\begin{aligned}\widetilde{S}_{12}(y^T) &= -\max\{K_2 - \Pi_2(y_2^T), 0\} + \widetilde{L}_{12}(y^T) + \frac{1}{n-1}D_1(y^T)\\ &\geq -\max\left\{\left|\frac{\underline{w}_{12}}{\overline{w}_{21}}\right|(\bar{\Pi}_1 - K_1), K_2 - \underline{\Pi}_2\right\} + \frac{1}{n-1}D_1(y^T).\end{aligned} \tag{49}$$

By our choice of $\overline{w}_{ij}$ and $\underline{w}_{ij}$, for some $\epsilon_1 > 0$,

$$\begin{aligned}\bar{\Pi}_1 - K_1 &\equiv \frac{1-\delta^T}{1-\delta}\Big(\max_{y_1} r_1(a_1^*, y_1) - g_1(a^*)\Big) - T^{2/3}\\ &< \frac{1-\delta^T}{1-\delta}\left(\frac{2}{n(n-1)}\overline{w}_{21} - \epsilon_1\right) - T^{2/3}\end{aligned} \tag{50}$$

and

$$\begin{aligned}K_2 - \underline{\Pi}_2 &\equiv \frac{1-\delta^T}{1-\delta}\Big(g_2(a^*) - \min_{y_2} r_2(a_2^*, y_2)\Big) + T^{2/3}\\ &< -\frac{1-\delta^T}{1-\delta}\left(\frac{2}{n(n-1)}\underline{w}_{12} + \epsilon_1\right) + T^{2/3}.\end{aligned} \tag{51}$$

Substituting (50) and (51) into (49), we have

$$\begin{aligned}\frac{(1-\delta)\delta^{-T}}{1-\delta^T}\widetilde{S}_{12}(y^T) \geq \frac{2}{n(n-1)}\underline{w}_{12} + (\delta^{-T} - 1)\frac{2}{n(n-1)}\underline{w}_{12} + \delta^{-T}\epsilon_1&\\ - \frac{(1-\delta)\delta^{-T}}{1-\delta^T}\left(T^{2/3} - \frac{1}{n-1}D_1(y^T)\right).&\end{aligned} \tag{52}$$

We can choose T_2 such that for all $T \geq T_2$ and $\delta \geq 1 - T^{-2}$, the sum of the last three terms on the right-hand side of (52) is positive. Inequality (46) follows from the same argument.

We now turn to (47) and (48). There are four cases to consider, depending on the values of $f_{12}(y^T)$ and $f_{21}(y^T)$.

Case 1: $f_{12}(y^T) = f_{21}(y^T) = 0$. In this case, inequalities (47) and (48) hold as $\widetilde{S}_{12}(y^T)$ and $\widetilde{S}_{21}(y^T)$ are both negative.

Case 2: $f_{12}(y^T) = 1$, $f_{21}(y^T) = 0$. In this case, following the definitions of $\widetilde{S}_{12}$ and $\widetilde{S}_{21}$,

$$\begin{aligned}\widetilde{S}_{12}(y^T) &\leq -\max\{K_2 - \Pi_2(y_2^T), 0\} + \max\{\Pi_2(y_2^T) - K_2, 0\}\\ &\leq \max\{\Pi_2(y_2^T) - K_2, 0\}\end{aligned}$$

and

$$\begin{aligned}\widetilde{S}_{21}(y^T) &\leq -\max\{K_1 - \Pi_1(y_1^T), 0\} - \max\left\{\left|\frac{\underline{w}_{21}}{\overline{w}_{12}}\right|(\bar{\Pi}_2 - K_2) - \max\{K_1 - \Pi_1(y_1^T), 0\}, 0\right\}\\ &\leq \frac{\underline{w}_{21}}{\overline{w}_{12}}(\bar{\Pi}_2 - K_2).\end{aligned}$$

Hence,

$$\widetilde{S}_{12}(y^T) - \frac{\overline{w}_{12}}{\underline{w}_{21}}\widetilde{S}_{21}(y^T) \le \max\{\Pi_2(y_2^T) - K_2, 0\} - (\bar{\Pi}_2 - K_2)$$
$$\le 0$$

and

$$-\frac{\overline{w}_{21}}{\underline{w}_{12}}\widetilde{S}_{12}(y^T) + \widetilde{S}_{21}(y^T) \le -\frac{\overline{w}_{21}}{\underline{w}_{12}} \max\{\Pi_2(y_2^T) - K_2, 0\} + \frac{\underline{w}_{21}}{\overline{\overline{w}}_{12}}(\bar{\Pi}_2 - K_2)$$
$$\le 0.$$

Case 3: $f_{12}(y^T) = 0$, $f_{21}(y^T) = 1$. The proof for this case is similar to that of Case 2 and, hence, is omitted.

Case 4: $f_{12}(y^T) = f_{21}(y^T) = 1$. In this case,

$$\widetilde{S}_{12}(y^T) \le \max\{\Pi_2(y_2^T) - K_2, 0\} - \left|\frac{\underline{w}_{12}}{\overline{\overline{w}}_{21}}\right|(\bar{\Pi}_1 - K_1)$$
$$\widetilde{S}_{21}(y^T) \le \max\{\Pi_1(y_1^T) - K_1, 0\} - \left|\frac{\underline{w}_{21}}{\overline{\overline{w}}_{12}}\right|(\bar{\Pi}_2 - K_2).$$

Inequality (47) holds as

$$\begin{aligned}&\widetilde{S}_{12}(y^T) - \frac{\overline{w}_{12}}{\underline{w}_{21}}\widetilde{S}_{21}(y^T)\\ &\le \max\{\Pi_2(y_2^T) - K_2, 0\} - \left|\frac{\underline{w}_{12}}{\overline{\overline{w}}_{21}}\right|(\bar{\Pi}_1 - K_1) + \left|\frac{\overline{w}_{12}}{\underline{w}_{21}}\right|\max\{\Pi_1(y_1^T) - K_1, 0\} - (\bar{\Pi}_2 - K_2)\\ &\le 0.\end{aligned}$$

Inequality (48) follows from the same argument. □

Finally, we show that (41) holds when T is sufficiently large and δ is sufficiently close to 1.

Lemma 6. *There exists T_3 such that (41) holds for any $T \ge T_3$, $\delta \ge 1 - T^{-2}$, $a \in \mathcal{B}$, and $y^T \in Y^T$.*

Proof. Let $\mathcal{C}$ denote the set of distinct pairs of players $\{i, j\}$. For any $\{i, j\} \in \mathcal{C}$, define for each $y^T \in Y^T$,

$$\widetilde{S}^{a,\{i,j\}}(y^T) \equiv (\widetilde{S}_1^{a,\{i,j\}}(y^T), \dots, \widetilde{S}_n^{a,\{i,j\}}(y^T)),$$

where

$$\widetilde{S}_k^{a,\{i,j\}}(y^T) \equiv \begin{cases} \widetilde{S}_{kj}^a(y^T) & \text{if } k = i \\ \widetilde{S}_{ki}^a(y^T) & \text{if } k = j \\ 0 & \text{if } k \ne i, j. \end{cases}$$

It is straightforward to check that

$$\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}^a(y^T) = \sum_{\{i,j\}\in\mathcal{C}} \frac{2}{n(n-1)}\left(\frac{n(n-1)}{2}\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}^{a,\{i,j\}}(y^T)\right).$$

Hence,

$$\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}^a(y^T)$$

is a convex combination of

$$\left\{\frac{n(n-1)}{2}\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}^{a,\{i,j\}}(y^T)\right\}_{\{i,j\}\in\mathcal{C}}.$$

Recall that the points

$$(\overline{w}^a_{ij}, \underline{w}^a_{ji}), \qquad (\underline{w}^a_{ij}, \overline{w}^a_{ji}), \qquad (\underline{w}^a_{ij}, \underline{w}^a_{ji})$$

are chosen to be in the interior of $W(\mathcal{B})$. Hence, by Lemmas 4 and 5, we can choose T_3 such that for any $T \geq T_3$, $\delta \geq 1 - T^{-2}$, $a \in \mathcal{B}$, $\{i, j\} \in \mathcal{C}$, and $y^T \in Y^T$,

$$\frac{n(n-1)}{2}\frac{1-\delta}{1-\delta^T}\delta^{-T}\widetilde{S}^{a,\{i,j\}}(y^T) \in W^{T,\delta}(\mathcal{B}). \qquad \square$$

Given any $\epsilon > 0$, let $T_4 = \max\{T_0, T_3\}$ and $\bar{\delta} = 1 - T_4^{-2}$. Then Lemmas 3 and 6 hold for $T = T_4$ and $\delta \geq \bar{\delta}$. By Lemma 6, we can choose transition probabilities $\{\mu^a_l\}_{l\in\mathcal{B}\cup\{N\}}$ such that (18) holds for all $a \in \mathcal{B}$ and $y^T \in Y^T$. It follows from Lemma 3 that the trigger-strategy profile that we described in Section 5 constitutes a perfect T-public equilibrium in which the average equilibrium payoff for each player i is at least $g_i(a^*) - \epsilon$.

References

Abreu, Dilip, Paul Milgrom, and David Pearce (1991), "Information and timing in repeated partnerships." *Econometrica*, 59, 1713–1733. [55, 73]

Aoyagi, Masaki (2002), "Collusion in dynamic Bertrand oligopoly with correlated private signals and communication." *Journal of Economic Theory*, 102, 229–248. [56]

Athey, Susan and Ilya Segal (2013), "An efficient dynamic mechanism." *Econometrica*, 81, 2463–2485. [55]

Aumann, Robert J. and Michael Maschler (1995), *Repeated Games With Incomplete Information*. MIT Press, Cambridge, Massachusetts. [84]

Aumann, Robert J., Michael B. Maschler, and Richard E. Stearns (1968), "Repeated games with incomplete information: An approach to the non-zero-sum case." Report to the U.S. Arms Control and Disarmament Agency ST-140, Princeton, New Jersey. Published as Chapter V in Aumann and Maschler (1995). [56]

Ben-Porath, Elchanan and Michael Kahneman (1996), "Communication in repeated games with private monitoring." *Journal of Economic Theory*, 70, 281–297. [56]

Bergemann, Dirk and Juuso Välimäki (2010), "The dynamic pivot mechanism." *Econometrica*, 78, 771–789. [55]

Bhaskar, V. and Ichiro Obara (2002), "Belief-based equilibria in the repeated prisoners' dilemma with private monitoring." *Journal of Economic Theory*, 102, 40–69. [56]

Bhaskar, V. and Eric van Damme (2002), "Moral hazard and private monitoring." *Journal of Economic Theory*, 102, 16–39. [56]

Chan, Jimmy and Wenzhang Zhang (2012), "Collusion enforcement with private information and private monitoring." Working paper, Shanghai University of Finance and Economics. [54]

Compte, Olivier (1998), "Communication in repeated games with imperfect private monitoring." *Econometrica*, 66, 597–626. [54, 55, 57, 60]

Ely, Jeffrey C. and Juuso Välimäki (2002), "A robust folk theorem for the prisoner's dilemma." *Journal of Economic Theory*, 102, 84–105. [56]

Escobar, Juan F. and Juuso Toikka (2013), "Efficiency in games with Markovian private information." *Econometrica*, 81, 1887–1934. [55]

Fong, Kyna, Olivier Gossner, Johannes Hörner, and Yuliy Sannikov (2011), "Efficiency in a repeated prisoner's dilemma with imperfect private monitoring." Working paper, Princeton University. [54, 56, 60]

Fuchs, William (2007), "Contracting with repeated moral hazard and private evaluations." *American Economic Review*, 97, 1432–1448. [54]

Fudenberg, Drew, David Levine, and Eric Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62, 997–1039. [54, 55]

Fudenberg, Drew and David K. Levine (1994), "Efficiency and observability with long-run and short-run players." *Journal of Economic Theory*, 62, 103–135. [67, 73]

Fudenberg, Drew and David K. Levine (2007), "The Nash-threats folk theorem with communication and approximate common knowledge in two player games." *Journal of Economic Theory*, 132, 461–473. [56]

Harrington, Joseph E., Jr. (2006), "How do cartels operate?" *Foundations and Trends in Microeconomics*, 2, 1–105. [54]

Harrington, Joseph E., Jr. and Andrzej Skrzypacz (2007), "Collusion under monitoring of sales." *RAND Journal of Economics*, 38, 314–331. [54]

Harrington, Joseph E. and Andrzej Skrzypacz (2011), "Private monitoring and communication in cartels: Explaining recent collusive practices." *American Economic Review*, 101, 2425–2449. [54]

Hoeffding, Wassily (1963), "Probability inequalities for sums of bounded random variables." *Journal of the American Statistical Association*, 58, 13–30. [60]

Hörner, Johannes and Wojciech Olszewski (2006), "The folk theorem for games with private almost-perfect monitoring." *Econometrica*, 74, 1499–1544. [56]

Hörner, Johannes and Wojciech Olszewski (2009), "How robust is the folk theorem?" *Quarterly Journal of Economics*, 124, 1773–1814. [56]

Hörner, Johannes, Satoru Takahashi, and Nicolas Vieille (2013), "Truthful equilibria in dynamic Bayesian games." Discussion Paper 1933, Cowles Foundation for Research in Economics, Yale University. [55]

Jackson, Matthew O. and Hugo F. Sonnenschein (2007), "Overcoming incentive constraints by linking decisions." *Econometrica*, 75, 241–257. [55]

Kandori, Michihiro (2002), "Introduction to repeated games with private monitoring." *Journal of Economic Theory*, 102, 1–15. [56]

Kandori, Michihiro and Hitoshi Matsushima (1998), "Private observation, communication and collusion." *Econometrica*, 66, 627–652. [55, 58]

Levin, Jonathan (2003), "Relational incentive contracts." *American Economic Review*, 93, 835–857. [54]

Mailath, George J. and Stephen Morris (2002), "Repeated games with almost-public monitoring." *Journal of Economic Theory*, 102, 189–228. [56]

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, Oxford. [56]

Matsushima, Hitoshi (2004), "Repeated games with private monitoring: Two players." *Econometrica*, 72, 823–852. [56]

Nissim, Kobbi, Rann Smorodinsky, and Moshe Tennenholtz (2012), "Approximately optimal mechanism design via differential privacy." Technical report, Israel Institute of Technology. [54]

Obara, Ichiro (2009), "Folk theorem with communication." *Journal of Economic Theory*, 144, 120–134. [54, 55, 57, 59]

Piccione, Michele (2002), "The repeated prisoner's dilemma with imperfect private monitoring." *Journal of Economic Theory*, 102, 70–83. [56]

Radner, Roy (1985), "Repeated principal–agent games with discounting." *Econometrica*, 53, 1173–1198. [55]

Radner, Roy, Roger Myerson, and Eric Maskin (1986), "An example of a repeated partnership game with discounting and with uniformly inefficient equilibria." *Review of Economic Studies*, 53, 59–69. [73]

Rahman, David (2012), "But who will monitor the monitor?" *American Economic Review*, 102, 2767–2797. [59]

Rubinstein, Ariel and Menahem E. Yaari (1983), "Repeated insurance contracts and moral hazard." *Journal of Economic Theory*, 30, 74–97. [55]

Sekiguchi, Tadashi (1997), "Efficiency in repeated prisoner's dilemma with private monitoring." *Journal of Economic Theory*, 76, 345–361. [56]

Stigler, George J. (1964), "A theory of oligopoly." *Journal of Political Economy*, 72, 44–61. [53]

Sugaya, Takuo (2012), "Folk theorem in repeated games with private monitoring." Working paper, Princeton University. [56]

Yamamoto, Yuichi (2007), "Efficiency results in N player games with imperfect private monitoring." *Journal of Economic Theory*, 135, 382–413. [56]

Zheng, Bingyong (2008), "Approximate efficiency in repeated games with correlated private signals." *Games and Economic Behavior*, 63, 406–416. [55]

Submitted 2012-10-19. Final version accepted 2014-11-30. Available online 2014-12-1.