The Value of Linking: Efficiency and Public Punishment

Jimmy Chan† Wenzhang Zhang‡

November 2, 2015

Abstract

We introduce a notion of public punishment to characterize the value of linking in repeated games with communication and side-payments. For any punishment scheme that enforces a stage-game outcome, the amount of public punishment is the expected payoff difference across self-evident sets. If a stage-game action profile with total stage-game payoff $g$ can be strictly enforced with total public punishment $L$, then for any $\epsilon > 0$ there exists an equilibrium in the repeated game with total average payoff greater than $g - L - \epsilon$ as the discount factor goes to one. If the signal distribution has full support, any strictly enforceable outcome can be enforced with no public punishment. For any strictly enforceable outcome, there is a $\epsilon$-close correlated outcome that is strictly enforceable with no public punishment. Hence, approximate efficiency can always be attained when an efficient outcome is strictly enforceable and the players can play correlated strategies.

1 Introduction

In recent years the theory of repeated games has made substantial progress in understanding how disperse, noisy information can be utilized to enforce cooperative

---

*Preliminary and Incomplete.
†Fudan University. Email: jimmy.hing.chan@gmail.com
‡Shanghai University of Finance and Economics. Email:wenzhang.zhang@gmail.com
behavior in long-run relationships. An important insight that has emerged from this literature is that linking punishment decisions across periods may reduce the cost of imperfect monitoring.

In a seminal paper, Abreu, Milgrom, and Pearce (1991) consider a two-person repeated Prisoners’ Dilemma. To induce cooperation both players must be punished at a cost when a bad signal occurs. Abreu, Milgrom, and Pearce (1991) show that if the signals are publicly observed immediately at the end of each period, the costly punishment must be carried out. But if the signals are observed with a lag, the players can delay the punishment and use the same punishment to induce cooperation in multiple periods. Intuitively, not observing the signals immediately reduces the number of incentive compatibility constraints. Subsequent research has shown that in repeated games of private monitoring a similar result may be obtained when players delay revealing their private signals endogenously.

In this paper we extend the insights of Abreu, Milgrom, and Pearce (1991) to a more general environment. We relate the value of linking to a notion of public punishment and introduce a new approach to attain approximate efficiency through correlated strategies. In our model players may observe private signals, play correlated strategies through a mediator, and exchange side-payments. There is no restriction on the signal distribution. At the end of each period each player form a belief over the actions and signals of the other players on the basis of his own. A set of action-signal profiles is self-evident if whenever the true action-signal profile is in the set, every player knows that it is in the set, and every player knows that every player knows that it is in the set, and so on. In equilibrium any difference in expected continuation payoffs across self-evident sets must result in efficiency loss because it is publicly known. More interestingly, any “non-public” punishment—payoff difference across action-signal profiles within a non-reducible self-evident set can be eliminated through linking. Specifically, we show that if an action profile which results in total stage-game payoff $g$ can be strictly enforced

\footnote{While uncommon in repeated games, side-payments are natural in many economic problems. Allowing side-payments allow us to avoid the difficulty of enforcing an outcome with rewards (rather than punishments).}
with total public punishment $L$, then for any $\epsilon > 0$ there exists an equilibrium in the repeated game with total average payoff greater than $g - L - \epsilon$ as the discount factor goes to one.

When the signal distribution has full support, no player can use his own signal to rule out any signal of another player. In that case there is only one self-evident set and, by definition, there is always no public punishment. It follows that with full support approximate efficiency is attainable if an efficient action is strictly enforceable. Furthermore, for any strictly enforceable efficient outcome, there always exists a $\epsilon$-close correlated outcome that can be strictly enforced without public punishment. For example, in the Prisoners’ Dilemma in Abreu, Milgrom, and Pearce (1991), the mediator may secretly tell one of the players to deviate with a very small probability. When a player is told to deviate, he will be rewarded when the outcome is a bad signal. Now conditional on a bad signal neither player can be sure that both of them will be punished when a bad signal occurs. Since deviating maximizes the chance of a bad signal, it is optimal for a player to deviate when told to do so. We say that asking a player to deviate provides a “benchmark” that prevents the punishment to deter the player from deviating when told to cooperate to become public. We show that in general we can always find enough benchmarks so that no punishment is public. Since the chance of being asked to choose a benchmark can be made very small, the efficiency loss from choosing the benchmarks is negligible. This result implies when the players can play correlated strategies, a sufficient condition for approximate efficiency is that an efficient outcome is strictly enforceable. In two-player games a mediator is needed to implement a correlated strategy. With more than three players, correlation can be achieved through with private communication between the players.

Our results are related to the literature that extends Abreu, Milgrom, and Pearce (1991) to repeated games with private monitoring (Compte (1998), Obara (2009), and Chan and Zhang (forthcoming)). In these papers players link the punishment decisions within a $T$-period block by delaying the report of private signals within a $T$-period block till the end of the block. Compared to Abreu, Milgrom, and Pearce (1991), the main difference is that with private monitoring
each player will update his beliefs about the outcomes during a $T$-period block through his own signals. Compte (1998) avoids this issue by assuming that players’ signals are independent. Obara (2009) considers correlated signals and identifies a condition on the signal distribution that ensures that no player can learn about his own continuation payoff. Chan and Zhang (forthcoming) introduce a different approach. Instead of preventing players from learning their own punishments, they make other players pay subsidies to players who learn “too much”. Chan and Zhang (forthcoming) assume that the signal distribution has full support. We generalize the results of Chan and Zhang (forthcoming) and show that what really matters is the expected payoff difference across self-evident sets. Our extension allows us to apply the idea of public punishment to games with any monitoring structures and correlated strategies.

The value of correlated strategies has been demonstrated by Rahman (2012). He shows that players may use a correlated strategy to achieve approximate efficiency when the signals are public. In Rahman (2012) conditional on the same signal each player is rewarded differently depending the actions of the other players. Correlated actions thus serve as a “secret test” that prevents a player from learn about his own punishment.\textsuperscript{2} By contrast, in our model conditional on the same signal each player is rewarded differently depending the action the player himself is told to choose by the mediator. The purpose is not to prevent one player from learning the punishment of another player. Rahman (2012) approach requires that an efficient outcome be identifiable conditional on every realization of the public signal. Ours requires only that an efficient action be (unconditionally) strictly identifiable. When each player can observe his own payoff, a strictly efficient outcome is always strictly identifiable.

\textsuperscript{2}Thus, Rahman (2012) is similar to Obara (2009) in that both try to prevent a player from learning anything about his own punishment.
2 Model

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

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

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

Let $a_{-i}$ and $y_{-i}$ denote $a$ minus $a_i$ and $y$ minus $y_i$, respectively. The marginal probabilities of $y_{-i}$, $y_i$ and $(y_i, y_j)$ are denoted respectively by $p_{-i}(y_{-i}|a)$, $p_i (y_i|a)$ and $p_{ij}(y_i, y_j|a)$, and the marginal probabilities of $y_{-i}$ and $y_j$, conditional on $a$ and $y_i$, are denoted respectively by $p_{-i}(y_{-i}|a, y_i)$ and $p_j (y_j|a, y_i)$. To simplify exposition, we assume that the support of $p$ is independent of $a$ and use $Y^*$ to

---

3We include $\tau_{ii}$, player $i$’s payment to himself, to simplify notation. Throughout, we set $\tau_{ii}$ to zero.

4For any variable $x_i$ we use $x$ to denote $(x_1, x_2, \ldots, x_n)$ and $x_{-i}$ to denote $x$ with the $i$-th element $x_i$ deleted.
denote the common support. Beyond common support we do not impose any other restrictions on \( p \).

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

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

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

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

The mediator is indifferent over all outcomes.

For each variable \( x \), we use \( x_t \) to denote the period- \( t \) value of \( x \), and \( x^t \) to denote the history of \( x \) up to period \( t \). For example, \( m^t = (m_0, m_1, \ldots, m_t) \) is the history of players’ message-report profiles up to period \( t \). We use \( x_{i,t} \) to denote period- \( t \) value of \( x_i \). Because it will usually be clear from the context that whether, say, \( a_1 \), refers to the action of player 1 or the action profile chosen in period 1, we do not introduce extra notations to distinguish the two. We will state the meaning explicitly when confusion may arise.

A history of the game is an infinite sequence

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

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

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

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

\[
u_{i,s} (h^\infty) \equiv (1 - \delta) \sum_{t=s}^{\infty} \delta^t \left( r_i(a_{i,t}, y_{i,t}) + \sum_{j=1}^{n} (\tau_{ji,t} - \tau_{ij,t}) \right).
\]
At the beginning of each period $t$, both the mediator and all players have observed a public history $h_{i,t}^{pub}$ that consists of the signal reports of the players, public message of the mediator, side-payments, and outcomes of the public randomization device in the previous $t$ periods. Let $H_t^{pub}$ denote the set of period-$t$ public histories, and let

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

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

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

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

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

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

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

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

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

$$H_{i,t} = H_{i,t}^1 \cup H_{i,t}^2 \cup H_{i,t}^3$$
denote the set of possible histories for player $i$ during period $t$ after which he needs
to make a decision. A generic element of $H_{i,t}$ will be denoted by $\varphi_{i,t}$.

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

A system of beliefs $\mu$ specifies, for each $i \in N$ and each $t \geq 0$, a probability
distribution $\mu_{\varphi_{i,t}}(\cdot)$ for each $\varphi_{i,t} \in H_{1,i,t}$, and a probability
distribution $\mu_{\varphi_{i,t}}(\cdot)$ for each $\varphi_{i,t} \in H_{2,i,t} \cup H_{3,i,t}$. An assessment
$(\sigma, \mu)$ consists of a strategy profile and a system of beliefs. Given any assessment
$(\sigma, \mu)$ and any history $\varphi_{i,t} \in H_{i,t}$, we use
\[
v_{i,t}(\sigma, \mu, \varphi_{i,t}) \equiv E\left[u_{i,t}(h^\infty) \mid \sigma, \mu, \varphi_{i,t}\right]
\]
to denote the expected value of player $i$’s payoff from period $t$ onward, where the
expectation is taken over the distribution of histories $h^\infty$ induced by $\sigma$, conditional
on $\varphi_{i,t}$ and the belief $\mu_{\varphi_{i,t}}(\cdot)$. Write $v_i(\sigma, \mu)$ for $v_{i,0}(\sigma, \mu, \varphi_{i,0})$.

An assessment $(\sigma_m, \sigma, \mu)$ is a perfect Bayesian equilibrium if the following two
conditions hold.

- For all $i \in N$, $t \geq 0$, and $\varphi_{i,t} \in H_{i,t}$ such that $\Pr(\varphi_{i,t} \mid \sigma) > 0$, the belief
  $\mu_{\varphi_{i,t}}(\cdot)$ is derived from $(\sigma_m, \sigma)$ using Bayes’ rule.

- For all $i \in N$, $t \geq 0$, $\varphi_{i,t} \in H_{i,t}$, and $\sigma'_i \in \Sigma_i$,
  \[
v_{i,t}(\sigma, \mu, \varphi_{i,t}) \geq v_{i,t}(\sigma'_i, \sigma_{-i}, \mu, \varphi_{i,t}) \tag{1}
  \]

Note that for $\varphi_{i,t} \in H_{i,t}^2 \cup H_{i,t}^3$, (1) is equivalent to
\[
E \left[ (1 - \delta)\delta^t \sum_{j=1}^n (\tau_{ji,t} - \tau_{ij,t}) + u_{i,t+1}(h^\infty) \right] \mid \sigma, \mu, \varphi_{i,t} \right] \geq E \left[ (1 - \delta)\delta^t \sum_{j=1}^n (\tau_{ji,t} - \tau_{ij,t}) + u_{i,t+1}(h^\infty) \mid \sigma'_i, \sigma_{-i}, \mu, \varphi_{i,t} \right].
\]
Hence the second condition implies sequential rationality.
3 Self-Evident Events

A key question of our analysis is to define what is public when the signals are private. Fix the mediator strategy \( \eta \) and assume the players follow the recommendations of the mediator. Let

\[
\mu (\tilde{a}, y) \equiv p(y|\tilde{a}) \eta (\tilde{a})
\]

be the distribution of \((\tilde{a}, y)\) induced by \( \eta \) and \( p \), and

\[
( A \times Y ) (\eta) \equiv \{ (\tilde{a}, y) | \mu (\tilde{a}, y) > 0 \}
\]

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

\[
P_i (\tilde{a}_i', y_i') = \{ (\tilde{a}_i, \tilde{a}_{-i}', y_i, y_{-i}) \in (A \times Y) (\eta) \}
\]

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

We can think of each subset \( E \) of \((A \times Y) (\eta)\) is an event. In the terminology of interactive epistemology, player \( i \) “knows” an event when he observes \((\tilde{a}_i, y_i)\) if \( P_i (\tilde{a}_i, y_i) \subseteq E \) in the sense that he knows the realized \((\tilde{a}, y)\) must belong to \( E \). An event \( E \) is common belief at \((\tilde{a}, y)\) if every player \( i \) knows \( E \), knows every player \( j \neq i \) knows \( E \), and so on when \((\tilde{a}, y)\) occurs. An event \( E \) is self-evident if it is common belief at any \((\tilde{a}, y) \in E \). Thus, \( E \) is self-evident if it is common belief whenever it occurs. Note that when a signal is public, its realization must be self-evident. In the following we argue that the converse is also true in our model. Each self-evident event is a “public” because each player \( i \) knows that the event has occurred, and each player knows that the other players knows the event has occurred, and so on.

Let \( P \) denote the meet (i.e., the least common coarsening) of \( \{P_i\}_{i=1}^n \). It is well known that any element of \( P \) is self-evident and any proper subset of any element of \( P \) is not (Osborne and Rubinstein Ch.5).\(^5\) The structure of self-evident sets plays an important role in the subsequent analysis.

\(^5\)Formally, for any \( \omega \in P, i \in \mathcal{N} \) and \((\tilde{a}_i', y_i') \in (A_i \times Y_i) (\eta)\), \( \mu (\cdot | (\tilde{a}_i', y_i'), \omega) = \mu (\cdot | (\tilde{a}_i', y_i')). \)
4 Public Punishment and Maximum Equilibrium Payoff

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

A pure strategy for player \( i \) in this extended game consists of an action strategy \( \alpha_i \) that maps each private recommendation into an action, and a reporting strategy \( \rho_i \) that maps each \( y_i \) into a message in \( Y_i \). Let \( \Psi_i \) the set of reporting strategies.

We say that a punishment scheme \( w = (w_1, \ldots, w_n) \) enforces \( \eta \) if it is a Nash equilibrium for the players to follow the recommendations and report their signal truthfully when the mediator selects \( \tilde{a} \) according to \( \eta \).

**Definition 1.** A transfer function profile \( w = (w_1, \ldots, w_n) \) enforces \( \eta \) if, for each player \( i \) and each \( \tilde{a}_i \in \text{supp}(\eta_i) \), \( a'_i \in A_i \) and \( \rho'_i \in \Psi_i \),

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

(2)

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

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

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

Second, for any \( \omega \in P \), and \( x \subset \omega \), there exists \( i \in \mathcal{N} \) and \((\tilde{a}_i', y_i') \in (A_i \times Y_i)(\eta)\) such that \( \mu(x | (\tilde{a}_i', y_i')) > 0 \) and \( \mu(\cdot | (\tilde{a}_i', y_i'), x) \neq \mu(\cdot | (\tilde{a}_i', y_i')). \)
denote the total expected transfer.

For any $w \in \mathcal{W}(\eta)$ let

$$L(w, \eta) = W(w, \eta) - \max_{\omega \in P} \sum_{i} \sum_{\tilde{\alpha}, y} w_i(\tilde{\alpha}, y) \mu(\tilde{\alpha}, y|\omega).$$

Since at the end of each stage game every player knows which $\omega \in P$ has occurred, we can essentially think of the continuation game after each $\omega \in P$ as a separate subgame. Since incentives across self-evident sets must be separately provided, the best any linking mechanism can do is to reduce $\max_{\omega \in P} \sum_{i} \sum_{\tilde{\alpha}, y} w_i(\tilde{\alpha}, y) \mu(\tilde{\alpha}, y|\omega)$ to zero. Since $L(w, \eta)$ is the expected total transfers that varies across self-evident sets, we shall refer to $|L(w, \eta)|$ as the public punishment (associated with $w$).

**Theorem 1.** For any strictly enforceable $\eta$, $w \in \mathcal{W}(\eta)$, and $\epsilon > 0$, there exists $\delta < 1$ such that for each $\delta > \delta$, there is a perfect action-public equilibrium of $\Gamma^\infty(B, \chi, G, \delta)$ with total average equilibrium payoff greater than

$$\sum_{i=1}^{n} \sum_{a} g_i(\tilde{a}) \eta(\tilde{a}) + L(w, \eta) - \epsilon.$$

We have argued that public punishment cannot be reduced through linking. Theorem 1 shows that the converse is also true. If we can strictly enforce $\eta$ with punishment $|L(w, \eta)|$, then there is an equilibrium with total average payoff arbitrarily close to the total stage-game payoff when $\eta$ is played, minus the public punishment. The key to Theorem 1 is a new $T$-period linking mechanism that exploits the heterogenous beliefs between the players to ensure that players have the right incentives in all $T$ periods. With the mechanism, it is straightforward to construct the required perfect action-public equilibrium along the lines of Chan and Zhang (forthcoming).

To illustrate the idea behind the linking mechanism, consider two-person game with binary signals. After the players choose their actions, each player $i$ observes a private $y_i \in \{h_i, l_i\}$. The signal distribution conditional on $a^*$, the action profile to be enforced, is given in the following table.

<table>
<thead>
<tr>
<th></th>
<th>$h_2$</th>
<th>$l_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h_1$</td>
<td>2/5</td>
<td>2/5</td>
</tr>
<tr>
<td>$l_1$</td>
<td>0</td>
<td>1/5</td>
</tr>
</tbody>
</table>
Suppose player 1’s action can be strictly enforced by a transfer function

\[
    w_1(h_1, h_2) = \frac{5}{2};
\]

\[
    w_1(h_1, l_2) = w_1(l_1, l_2) = 0.
\]

If the stage game is repeated \( T \) times using this transfer function, then by the law of large number the total transfer will be unlikely to be significantly bigger than \( TE(w_1|a^*) = T \) when \( T \) is large. If player 1 did not observe his private signals until the end of all \( T \) periods (as in Abreu, Milgrom, and Pearce (1991)), then a nearly efficient way to induce player 1 to choose \( a^*_1 \) in all \( T \) periods would be to set his transfer at the end of the \( T \) periods to

\[
    \max \left( 0, TE(w_1|a^*) + \kappa - \sum_{t=1}^{T} w_1(y^t_1, y^t_2) \right),
\]

where \( y^t_i \) is player \( i \) in period \( t \) and \( \kappa \) is a small number. (For simplicity we assume \( \delta = 1 \) here.) Note that the transfer would be always negative, and the expected transfer would be close to \( \kappa \), which is small, when \( T \) is large. Since it is extremely unlikely that \( \sum_{t=1}^{T} w_1(y^t_1, y^t_2) \) is greater than \( TE(w_1|a^*) + \kappa \), player \( i \) is essentially facing a linear incentive scheme. Since the linear incentive scheme is strict, it would be optimal player \( i \) to choose \( a^*_1 \) in all \( T \) periods.

This simple scheme, however, does not work when player 1 observes \( y_1 \) at the end of each period. Note that \( E(w_1|a^*, h_1) = 5/4 \). If player 1 observes a large number of \( h_1 \) in the early periods, then he would come to believe that \( \sum_{t=1}^{T} w_1(y^t_1, y^t_2) \) is likely to be greater than \( TE(w_1|a^*) + \kappa \), and he would have no incentive to choose \( a^*_1 \) in the remaining periods. One way to deal with this “learning” problem is to change \( w_1 \) so that player 1’s expectation on \( w_1 \) does not change with \( y_i \). For example Compte (1998), Obara (2009). This approach, however, is not always feasible. In this example, because player 1 knows that \( y_2 = l_2 \) when he observes \( l_1 \), he would always learn about his only transfer if enforcing \( a^*_1 \) requires a transfer function that varies with \( h_2 \).

Our approach instead relies on the heterogeneous beliefs between players—although the players’ signals are correlated, player 2 does not know exactly what
player 1 knows. Define

\[ z_1(h_1, h_2) = 1, \quad z_1(h_1, l_2) = \frac{3}{2}, \quad z_1(l_1, l_2) = 0. \]

The function \( z_1 \) is constructed with the properties that

\[
E[z_1|a^*, h_1] = E[w_1|a^*, h_1] = \frac{5}{4};
\]

\[
E[z_1|a^*, l_1] = E[w_1|a^*, l_1] = 0;
\]

\[
E[z_1|a^*, h_2] = E[z_1|a^*, l_2] = E[z|a^*] = 1.
\]

Under our approach, player 1, in addition to

\[
\max\left(0, TE(w_1|a^*) + \kappa - \sum_{t=1}^{T} w_1(y^1_t, y^2_t)\right),
\]

will receive an additional payment from player 2 equal to

\[
\max\left(0, \sum_{t=1}^{T} w_1(y^1_t, y^2_t) - TE(w_1|a^*) - \kappa\right)
\]

when

\[
\sum_{t=1}^{T} z_1(y^1_t, y^2_t) \geq TE[z_1|a^*] + \kappa - \iota
\]

for some small \( \iota \in (0, \kappa) \).

To see how this scheme works, first consider the extreme case where player 1 has observed \( h_1 \) in each of the first \( T - 1 \) periods. When \( T \) is large, he will be pretty sure that player 2 has received about \( (T - 1)/2 \) \( h_1 \), which means that \( \sum_{t=1}^{T} w_1(y^1_t, y^2_t) \) will be greater than \( TE(w_1|a^*) + \kappa \). However, in this case \( \sum_{t=1}^{T} z_1(y^1_t, y^2_t) \) would also be greater \( TE[z_1|a^*] + \kappa - \iota \) because \( z_1 \) and \( w_1 \) has the same mean conditional on \( h_1 \), and \( \iota > 0 \) so that it is easier for \( \sum_{t=1}^{T} z_1(y^1_t, y^2_t) \) to cross the threshold and, hence, player 1 will receive the additional payment, which exactly makes up for the missing incentives. As a result, a player 1 who has \( T - 1 \) \( h_1 \) will believe that he is essentially facing a transfer function \( w_1 \) in the last period. It is thus optimal for him to choose \( a^*_1 \) in the last period. Note that, \( \sum_{t=1}^{T-1} z_1(h_1, y^2_t) \) is decreasing in the number of \( h_2 \). Player 1, therefore, will not receive the extra compensation when
player 2 has observed all \( h_2 \) in the first \( T - 1 \) periods. It does not affect player 1’s incentive significantly because since \( h_1 \) and \( h_2 \) are not perfectly correlated, player 1 believes that it is extremely unlikely for player 2 to observe all \( h_2 \) even when player 1 observes all \( h_1 \). Similar logic applies throughout \( T \) periods and regardless of the number of \( h_1 \) signals player 1 has observed—under this transfer scheme player 1 will always believe that he will receive the extra payment when he believes that there is non-trivial chance that \( \sum_{t=1}^{T} w_1(y_1^t, y_2^t) \) will be greater than \( TE(w_1|a^*) + \kappa \).

Thus far, we have considered only the incentives of player 1. Making player 2 pays a payment to player 1 in general may distort player 2’s incentives. But under our construction the distorting is negligible when \( T \) is large. By construction of \( z_1 \) player 2 will have to pay only when the player 1 observes a large number of \( h_1 \) and player 2 observes a “right” balance of \( h_2 \) and \( l_2 \). From player 2’s perspective, this event is very unlikely when \( T \) is large. If he observes a large number of \( h_2 \), he knows that he would not need to pay because \( z_1 \) is small when \( y_2 = h_2 \). On the other hand, if he observes a large fraction of \( l_2 \), he believes that there is a large probability that player 1 has observed a fair number of \( l_1 \), and in that case no payment is required. Another way to see this is notice that player 2’s expectation of \( z_1 \) is always equal to the unconditional mean of \( z_1 \), independent of his own signal. Hence, when \( T \) is large, player 2 always expects \( \sum_{t=1}^{T} z_1(y_1^t, y_2^t) \) to be close to \( TE[z_1|a^*] \) and unlikely to be greater than \( TE[z_1|a^*] + \kappa - \iota \) (as \( \kappa - \iota > 0 \)).

Note that the scheme above critically depends on the fact the players’ signals are not perfectly correlated. If the signals were perfectly correlated, then player 2 will know perfectly when player 1 will need an extra payment, and the need to make this payment will distort player 2’s incentives. When the signals are perfectly correlated, then any punishment must be public in the sense that every player knows every knows that some player needs to be punished, and so on. In the formal proof of Theorem 1, we show that almost all non-public punishment can be eliminated when the players are sufficiently patient.
5 Approximate Efficiency

Theorem 1 relates the attainable average equilibrium payoff in a repeated game to the public punishment needed to enforce an action profile in the stage game. The following theorem shows that for any strictly enforceable $\eta$, there is a correlated action profile close to $\eta$ that is strictly enforceable with zero public punishment.

**Theorem 2.** If $\eta$ is strictly enforceable, then for any $\epsilon$ there exists $\tilde{\eta}$ such that

$$\max_{a \in A} \| \eta(a) - \tilde{\eta}(a) \| \leq \epsilon$$

and $L^*(\tilde{\eta}) = 0$.

An immediate consequence of Theorem 2 is that approximate is attainable if an efficient action profile is strictly virtually enforceable.

We prove Theorem 2 in two steps. Let $L^*(\eta) \equiv \sup_{w \in W^*(\eta)} L(w, \eta)$ denote the supremum of the total public transfer required to enforce $\eta$. We first apply the theory of alternatives to characterize the dual of $L^*(\eta)$ and then use the dual to prove Theorem 2.

Let

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

denote the set of distributions that have the same conditional distributions as $\mu$, and let $s_i : A_i \to \Delta(A_i \times \Psi_i)$ denote a mixed strategy in $G_0$ where $s_i(\bar{a}_i, a_i, \rho_i)$ is the probability of choosing $(a_i, \rho_i)$ after receiving the recommendation $\bar{a}_i$. The distribution of $(\bar{a}, y)$ induced by $(s_i, \alpha_{-i}^*)$ is $\pi^{s_i}$. For each $(\bar{a}, y) \in A \times Y$,

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

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

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

Let

$$\gamma(s_i) \equiv \left( \max_{\omega \in P} \frac{\tilde{\nu}(\omega)}{\mu(\omega)} - 1 \right)^{-1}.$$
Proposition 1. For any strictly enforceable $\eta$, $L^* (\eta)$ is equal to

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

s.t. $\pi^{s_1} = \cdots = \pi^{s_n} \in Q(\eta)$. \hspace{1cm} (3)

Proposition 2 says that public punishment is needed only if there exists a profile of deviating strategies that lead to the same distribution that has the same conditional distributions as $\mu$. If $\pi^{s_i} \notin Q(\eta)$, then $s_i$ can be deterred by using private signals within some $\omega \in P$. If there is no $s_j$ such that $\pi^{s_j} = \pi^{s_i}$, then player $j$ cannot cause $\pi^{s_i}$, and we can punish player $i$ and reward player $j$ when $\pi^{s_i}$ “occurs” to keep the total punishment zero. For any $s_1, \ldots, s_n$ that satisfy (2), the minimum total punishment needed to deter each player $i$ from choosing $\gamma(s_1) \left[ \sum_{i=1}^{n} \sum_{\tilde{a} \in A} \eta(\tilde{a}) \sum_{a_i, \rho_i} s_i(\tilde{a}_i, a_i, \rho_i) \left( g_i(\tilde{a}) - g_i(\tilde{a}_{-i}, a_i) \right) \right]$, where the term inside the square bracket is the total deviation gains, and $\gamma(s_1)$ is the efficacy of the best statistical test (i.e., the counterpart of $(1 - p) / (p - q)$). The total public punishment needed to enforce a distribution is the public punishment needed to deter the “worst” deviation.

As an example, consider a Noisy Prisoners’ Dilemma game, where each player chooses $C$ or $D$ and then observes a public signal that is either $H$ or $L$. The stage-game payoffs and signal distribution are given in Figure 1. It is assumed that $p > q$ and $h, d > 0$.

<table>
<thead>
<tr>
<th>Actions</th>
<th>Public Signal Dist.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$H$</td>
</tr>
<tr>
<td>$C$</td>
<td>1, 1</td>
</tr>
</tbody>
</table>

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

$$P_1(C, H) = \{H\}, P_1(C, L) = \{L\};$$
$$P_2(C, H) = \{H\}, P_2(C, L) = \{L\};$$
$$P = \{H\}, \{L\}.$$
Since each member of $P$ is a singleton, $Q(\eta)$ contains any distributions over $H$ and $L$. In this game, the only deviation is $D$. If player $i$ chooses $D$, the signal $H$ will occur with probability $q$ and

$$\gamma = \frac{1 - p}{p - q}.$$  

Each player can gain $d$ by deviating to $D$. Hence, $L^*(CC)$ is $2d(1 - p)/(p - q)$. The factor $(1 - p)/(p - q)$ measures the efficacy of the transfer scheme. This factor is smaller if the deviation leads to a larger change in probability distribution (i.e., a bigger $p - q$) or if the “bad” signal $L$ is unlikely to occur when the players do not deviate.

We prove Theorem 2 by constructing, for any strictly enforceable $\eta$, a $\tilde{\eta}$ that is $\epsilon$-close to $\eta$ and such that there is no $(s_1, ..., s_n) \in Q(\tilde{\eta})$. Talk about the implication of this proposition. To illustrate the idea, return to the noisy Prisoners’ Dilemma example. Let $p(H|DD) = r$. Consider the distribution

$$\tilde{\eta}(CC) = 1 - \epsilon; \tilde{\eta}(CD) = \tilde{\eta}(DC) = 0.5\epsilon.$$  

The meet induced by $\tilde{\eta}$ contains two elements, but now each element contains three members. In particular

$$P = \{CCH, CDH, DCH\}, \{CCL, CDL, DCL\}.$$  

Let $(xy)$ denote the strategy of choosing $x$ when recommended to choose $C$ and $y$ when recommended to choose $D$. Each player has four pure strategies: $CD$, $DD$, $CC$, and $DC$. In the following table each row corresponds to the probability of each outcome that involves a $H$ signal for each pure strategy of player 1 (assuming that player 2 does not deviate and plays $(CD)$).

<table>
<thead>
<tr>
<th></th>
<th>$CCH$</th>
<th>$DCH$</th>
<th>$CDH$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$CD$</td>
<td>$(1 - \epsilon)p$</td>
<td>$0.5\epsilon q$</td>
<td>$0.5\epsilon q$</td>
</tr>
<tr>
<td>$DD$</td>
<td>$(1 - \epsilon)q$</td>
<td>$0.5\epsilon q$</td>
<td>$0.5\epsilon r$</td>
</tr>
<tr>
<td>$CC$</td>
<td>$(1 - \epsilon)p$</td>
<td>$0.5\epsilon p$</td>
<td>$0.5\epsilon q$</td>
</tr>
<tr>
<td>$DC$</td>
<td>$(1 - \epsilon)q$</td>
<td>$0.5\epsilon p$</td>
<td>$0.5\epsilon r$</td>
</tr>
</tbody>
</table>
Notice that the ratio of $CCH$ over $DCH$ is strictly higher when player 1 follows the recommendation and plays $(CD)$. Hence, any deviation from $(CD)$ must lower the ratio of $CCH$ over $DCH$ conditional on $H$. It follows that there is no deviation $s_1$ with $\pi^{s_1} \in Q(\tilde{\eta})$. Intuitively, the recommendation $DC$ serves as a “benchmark” for player 1. Given that player 2 is following the recommendation, choosing $D$ minimizes the probability of $H$. There is no way player 1 that can further lower the probability of $H$ when the recommendation is $DC$. Hence, if player 1 deviates to $D$ when told to choose $C$, he must lower the ratio of $CCH$ over $DCH$.

6 Conclusion

We introduce a general notion of public punishment, characterize the minimum public punishment needed to enforce a strategy profile, and establish a sufficient condition for approximate efficiency. Our results offer a new perspective on existing results and lead to a new approach to establish approximate efficiency through correlated strategies.
A  Key Lemmas

In this section we prove a key lemma that will be needed in the proof of Theorem 1.

Lemma 1. For any player $i \in \mathcal{N}$ and any function $f : A \times Y \to R$, there exists a function $z_j : A \times Y \to R$ for each $j \neq i$ such that

1. \[ \sum_{(a_i,y_i)} \sum_{j \neq i} z_j(a,y) \Pr(a_{-i},y_{-i}|a_i,y_i) = (n-1) \sum_{(a_i,y_i)} f(a,y) \Pr(a_{-i},y_{-i}|a_i,y_i) \]
   for all $(a_i,y_i) \in A_i \times Y_i$;

2. \[ \sum_{(a_j, y_{-j})} z_j(a,y) \Pr(a_{-j},y_{-j}|a_j,y_j) = \sum_{(a,y)} f(a,y) \Pr(a,y|\omega), \text{ for all } j \neq i, \ \omega \in P, \text{ and } (a_j,y_j) \in A_j \times Y_j \text{ such that } P_j(a_j,y_j) \subseteq \omega. \]

Proof of Lemma 1. W.l.o.g., set $i = n$. Since it suffices to construct $f$ on each component $\omega \in P$ separately, we assume $P = \{Y^*_\}$. The two conditions can be written as

\[ \sum_{j \neq n} \sum_{(a_{-n},y_{-n})} z_j(a,y) \Pr(a,y) = (n-1) \sum_{(a_{-n},y_{-n})} f(a,y) \Pr(a,y) \]

for all $(a_{-n},y_{-n}) \in A_{-n} \times Y_{-n}$, and

\[ \sum_{(a_j, y_{-j})} z_j(a,y) \Pr(a,y) = \sum_{(a,y)} f(a,y) \Pr(a,y) \Pr(a_j,y_j) \]

for all $(a_j,y_j) \in A_j \times Y_j$.

Fix enumerations $(a^1, a^2, \ldots, a^{|A|})$ of $A$, $(y^1, y^2, \ldots, y^{|Y|})$ of $Y$, and, for each $j \in \mathcal{N}$, $(a_j^1, a_j^2, \ldots, a_j^{|A_j|})$ of $A_j$, and $(y_j^1, y_j^2, \ldots, y_j^{|Y_j|})$ of $Y_j$. For each $j$, let $B_j$ be the $(|A_j| \cdot |Y_j|) \times (|A| \cdot |Y|)$ matrix where, for each $k_j \leq |A_j|$, $l_j \leq |Y_j|$, $k \leq |A|$, and $l \leq |Y|$, the $(k_j \cdot |Y_j| + l_j, k \cdot |Y| + l)$-th element is $\Pr(a^k, y^l)$ if the $j$-th component of $a^k$ is $a_j^{k_j}$ and the $j$-th component of $y^l$ is $y_j^{l_j}$, and 0 otherwise. Thus, the only possible nonzero element of the $(k \cdot |Y| + l)$-th column is $\Pr(a^k, y^l)$ in the $(k_j \cdot |Y_j| + l_j)$-th row, where $k_j$ and $l_j$ are derived from $a_j^k = a_j^{k_j}$ and $y_j^l = y_j^{l_j}$.
Rewrite (5) and (6) as

\[
\begin{pmatrix}
B_1 & 0 & \ldots & 0 \\
0 & B_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & B_{n-1}
\end{pmatrix}
\begin{pmatrix}
z_1(a_1^1, y_1^1) \\
\vdots \\
z_1(a_1^{|A_1|}, y_1^{|Y_1|}) \\
\vdots \\
z_{n-1}(a_1^1, y_1^1) \\
\vdots \\
z_{n-1}(a_1^{|A_1|}, y_1^{|Y_1|})
\end{pmatrix}
= 
\sum_{(a,y)} f(a, y) \Pr(a, y) \Pr(a_1^1, y_1^1) \\
\vdots \\
\sum_{(a,y)} f(a, y) \Pr(a, y) \Pr(a_1^{|A_1|}, y_1^{|Y_1|}) \\
\vdots \\
(n-1) \sum_{(a_{-n}, y_{-n})} f(a_1^1, y_1^1, a_{-n}, y_{-n}) \Pr(a_1^1, y_1^1, a_{-n}, y_{-n}) \\
\vdots \\
(n-1) \sum_{(a_{-n}, y_{-n})} f(a_1^{|A_1|}, y_1^{|Y_1|}, a_{-n}, y_{-n}) \Pr(a_1^{|A_1|}, y_1^{|Y_1|}, a_{-n}, y_{-n})
\] .

(7)

By the Fredholm alternative, either (7) has a solution \(z_1, \ldots, z_{n-1}\), or there exists \(\lambda_{j,k} \in \mathbb{N}, k_j \leq |A_j|, l_j \leq |Y_j|\) for each \(j \in \mathcal{N}\) such that

\[
(\lambda_j, \ldots, \lambda_{j,k_j}y_{j_1}+l_j, \ldots, \lambda_{j,|A_j|}y_{j_1}) B_j + (\lambda_{n,1}, \ldots, \lambda_{n,k_n}y_{n_1}+l_n, \ldots, \lambda_{n,|A_n|}y_{n_1}) B_n = 0
\]

for each \(j \neq n\), and

\[
\sum_{j=1}^{n-1} \sum_{k_j=1}^{|A_j|} \sum_{l_j=1}^{|Y_j|} \lambda_{j,k_j}y_{j_1}+l_j \sum_{(a,y)} f(a, y) \Pr(a, y) \Pr(a_1^1, y_1^1) + \\
(n-1) \sum_{k_n=1}^{|A_n|} \sum_{l_n=1}^{|Y_n|} \lambda_{n,k_n}y_{n_1}+l_n \sum_{(a_{-n}, y_{-n})} f(a_1^1, y_1^1, a_{-n}, y_{-n}) \Pr(a_1^1, y_1^1, a_{-n}, y_{-n}) \neq 0.
\]

(9)

We show that (8) and (9) are incompatible. Hence, (7) has a solution.

Suppose that (8) holds. We want to show that for some \(\lambda\)

\[
\lambda_{n,k_n}y_{n_1}+l_n = \lambda \quad \text{for all } k_n \leq |A_n|, l_n \leq |Y_n|
\]

(10)
and
\[
\lambda_{j,k_j|Y_j|+I_j} = -\lambda \quad \text{for all } k_j \leq |A_j|, I_j \leq |Y_j| \text{ and all } j \neq n. 
\] (11)

Hence (9) is false. To prove (10), fix any \((k', l')\) and \((k'', l'')\) such that \(\Pr (a^{k'}, y^{l'}) > 0\) and \(\Pr (a^{k''}, y^{l''}) > 0\). By the definition of the meet \(\tilde{P}\) there exists a sequence of players \(i_1, i_2, \ldots, i_q\) and a sequence \((k', l') = (k(0), l(0)), (k(1), l(1)), \ldots, (k(q), l(q)) = (k'', l'')\) such that \((a^{k(s)}, y^{l(s)}) \in P_{i_s} (a^{k(s-1)}, y^{l(s-1)})\) for each \(s \leq q\).

Hence, for each \(s \leq q\), there exist \(\tilde{k}(s) \leq |A_{i_s}|\) and \(\tilde{l}(s) \leq |Y_{i_s}|\) such that the \(i_s\)-th components of \(a^{k(s)}\) and \(a^{k(s-1)}\) are both equal to \(a^{\tilde{k}(s)}\), and the \(i_s\)-th components of \(y^{l(s)}\) and \(y^{l(s-1)}\) are both equal to \(y^{\tilde{l}(s)}\). Let the \(n\)-th component of each \(a^{k(s)}\) and \(y^{l(s)}\) be \(a^{\tilde{k}(s)}_n\) and \(y^{\tilde{l}(s)}_n\), respectively. We show that

\[
\lambda_{n,\tilde{k}(0)|Y_n|+\tilde{l}(0)} = \lambda_{n,\tilde{k}(1)|Y_n|+\tilde{l}(1)} = \cdots = \lambda_{n,\tilde{k}(q)|Y_n|+\tilde{l}(q)}. \tag{12}
\]

For each \(s\), if \(i_s = n\), then \(\tilde{k}(s-1) = \tilde{k}(s) = \tilde{k}(s)\) and \(\tilde{l}(s-1) = \tilde{l}(s) = \tilde{l}(s)\). It follows that

\[
\lambda_{n,\tilde{k}(s-1)|Y_n|+\tilde{l}(s-1)} = \lambda_{n,\tilde{k}(s)|Y_n|+\tilde{l}(s)}.
\]

if \(i_s = j \neq n\), consider the \((k(s-1)|Y| + l(s-1))-th\) column of (8). Since the only possible nonzero elements of \(B_j\) and \(B_n\) in that column is \(\Pr (a^{s-1}, y^{s-1})\), in the \((\tilde{k}(s-1)|Y_n| + \tilde{l}(s-1))-th\) and \((\tilde{k}(s-1)|Y_n| + \tilde{l}(s-1))-th\) rows, respectively. Hence (8) implies that

\[
\lambda_{i_s,\tilde{k}(s)|Y_n|+\tilde{l}(s)} \Pr (a^{s-1}, y^{s-1}) + \lambda_{n,\tilde{k}(s-1)|Y_n|+\tilde{l}(s-1)} \Pr (a^{s-1}, y^{s-1}) = 0,
\]

or

\[
\lambda_{n,\tilde{k}(s-1)|Y_n|+\tilde{l}(s-1)} = -\lambda_{i_s,\tilde{k}(s)|Y_n|+\tilde{l}(s)}.
\]

Similarly,

\[
\lambda_{n,\tilde{k}(s)|Y_n|+\tilde{l}(s)} = -\lambda_{i_s,\tilde{k}(s)|Y_n|+\tilde{l}(s)}.
\]

This proves (12) \(\square\)

Lemma 2. Each function \(z_j\) in Lemma 1 can be constructed to satisfy the following additional conditions:

1. ...
Lemma 2. To show that it satisfies condition 1 of Lemma 1, fix any
by
\[ p_i (\cdot | a, y_i) = p_i (\cdot | a, \tilde{y}_i). \]

3. \[ z_j (a, y_j) = z_j (a, y_j, \tilde{y}_j) \text{ for all } y_j \in Y_j \text{ and } y_j, \tilde{y}_j \in Y_j \text{ such that} \]
\[ p_j (\cdot | a, y_j) = p_j (\cdot | a, \tilde{y}_j). \]

Proof of Lemma 2. Define a partition \( Y_j \) of \( Y_j \) for each \( j \) as follows. For any \( y_j, \tilde{y}_j \in Y_j \) and \( \eta_j \in Y_j \) such that \( y_j \in \eta_j \), we have \( \tilde{y}_j \in \eta_j \) if and only if \( p_j (\cdot | a, y_j) = p_j (\cdot | a, \tilde{y}_j) \). Define a probability distribution \( p \) on \( Y = Y_1 \times \cdots Y_n \) by
\[
p (\eta | a) = \prod_{i=1}^n p (y_i | a) \text{ for all } \eta.
\]

Since \( E [f(a, y) | a, y_i] \) is constant on each \( \eta_i \in Y_i \), it induces a function \( f : A \times Y_i \to R \). Apply Lemma 1 to the information structure \( (Y, p) \) and function \( f \), we obtain a function \( z_j : A \times Y \to R \) for each \( j \neq i \) such that
1. \[
\sum_{j \neq i} E [z_j (a, \eta) | a, \eta_i] = (n - 1) f (a, \eta_i) \text{ for all } \eta_i \in Y_i;
\]
2. \[
E [z_j (a, \eta) | a, \eta_i] = E [f (a, \eta_i) | a] \text{ for all } \eta_i \in Y_j.
\]

Now define \( z_j : Y \to R \) by letting \( z_j (a, y) = z_j (a, \eta) \) for each \( y \in Y \), where \( \eta \) is the element of \( Y \) that contains \( y \). By definition \( z_j \) satisfies conditions 1 and 2 of Lemma 2. To show that it satisfies condition 1 of Lemma 1, fix any \( \eta_i \in Y_i \) and \( y_i \in \eta_i \). We have
\[
\sum_{j \neq i} E [z_j (a, y) | a, y_i] = \sum_{j \neq i} \sum_{y_j \in Y_j} z_j (a, y) p_j (y_j | a, y_i)
\]
\[
= \sum_{j \neq i} \sum_{\eta_j \in Y_j} \sum_{y_j \in \eta_j} z_j (a, y) p_j (y_j | a, y_i)
\]
\[
= \sum_{j \neq i} \sum_{\eta_j \in Y_j} z_j (a, \eta) p_j (\eta_j | a, y_i)
\]
\[
= \sum_{j \neq i} \sum_{\eta_j \in Y_j} z_j (a, \eta) p_j (\eta_j | a, \eta_i)
\]
\[
= E [z_j (a, \eta) | a, \eta_i]
\]
\[
= (n - 1) f (a, \eta_i)
\]
\[
= (n - 1) E [f (a, y) | a, y_i].
\]
That $z_j$ satisfies condition 2 of Lemma 1 can be shown similarly.

B Proof of Theorem 1

The proof of Proposition Theorem 1 consists of two steps. First, we introduce an auxiliary $T$-period game in which the players play $G_0$ for $T$ periods, and report private signals and receive transfers at the end of period $T$. We design a profile of transfer functions to enforce $\eta$ in each period of this game with average efficiency loss less than $(L^*(\eta) + \epsilon)$. Next, we construct a $T$-period review-strategy perfect Bayesian equilibrium for the infinitely-repeated game by embedding the transfer functions for the $T$-period game into the repeated game.

B.1 Step 1: Enforcing $\eta$ in a $T$-period game

B.1.1 A $T$-period game

We consider the following $T$-period game, which we denote by $\Gamma^T(G_0, S, \delta)$. In each period $k = 1, \ldots, T$, the players play the stage-game $G_0$.

In each period $k$, the mediator always recommends $\tilde{a}(k)$ according to the distribution $\eta$. After receiving the recommendation $\tilde{a}_i(k)$, each player $i$ chooses a private action $a_i(k)$ and then receives the private signal $y_i(k)$. At the end of period $T$, each player $i$ simultaneously reports a message $m_i \in M_i$. In addition to his stage-game payoffs, each player $i$ receives a transfer $S_i(\tilde{a}^T, m)$ that depends on the recommended actions $\tilde{a}^T_i = (\tilde{a}_i(1), \ldots, \tilde{a}_i(T))$ and the profile of reports $m = (m_1, \ldots, m_n)$. Player $i$’s total discounted payoff in this $T$-period game is

$$\sum_{k=1}^{T} \delta^{k-1} r_i(a_i(k), y_i(k)) + S_i(\tilde{a}^T, m).$$

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

$$\sum_{i=1}^{n} S_i(\tilde{a}^T, m) \leq 0, \text{ for all } (\tilde{a}^T, m) \in A^T \times M. \quad (13)$$

$^6$Recall that $G_0$ denotes the stage game $G$ without the last three steps.
For any $k \leq T$, let $\tilde{a}_i^k \equiv (\tilde{a}_i(1), \ldots, \tilde{a}_i(k))$, $a_i^k \equiv (a_i(1), \ldots, a_i(k))$ and $y_i^k \equiv (y_i(1), \ldots, y_i(k))$ denote the recommended actions, private actions, and private signals of player $i$ in the first $k$ periods, respectively. Player $i$’s strategy consists of two components: an action strategy $\alpha_i^T$ that maps each $(\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1}) \in \cup_{l=1}^T (A_i^l \times A_i^{l-1} \times Y_i^{l-1})$ into an action $a_i \in A_i$ and a reporting strategy $\rho_i^T$ that maps each $(\tilde{a}_i^T, a_i^T, y_i^T) \in A_i^T \times A_i^T \times Y_i^T$ into a report $m_i \in M_i$.  

Let $A_i^T$ and $\Sigma_i^T$ be player $i$’s action-strategy set and reporting-strategy set, respectively. A system of beliefs $\mu^T$ specifies, for each player $i \in N$ and each history $(\tilde{a}_i^k, a_i^k, y_i^k) \in \cup_{i=0}^T (A_i^l \times A_i^l \times Y_i^l)$, a probability distribution $\mu_i^T (\cdot | \tilde{a}_i^k, a_i^k, y_i^k)$ of $(\tilde{a}_{i-1}^k, a_{i-1}^k, y_{i-1}^k)$. 

For each strategy profile $(\alpha^T, \rho^T)$, belief system $\mu^T$ and history $(\tilde{a}_i^k, a_i^k, y_i^k) \in \cup_{i=0}^T (A_i^l \times A_i^l \times Y_i^l)$, we let

$$
U_i^T (\alpha^T, \rho^T; S, \tilde{a}_i^k, a_i^k, y_i^k) = \mathbb{E}_{\tilde{a}_i^k, a_i^k, y_i^k} \left[ \sum_{l=1}^T \delta_l^{-1} g_i (a(l)) + S_i (\tilde{a}_i^T, \rho^T (\tilde{a}_i^T, a_i^T, y_i^T)) \bigg| \lambda_i^T, \mu^T, \alpha^T, \tilde{a}_i^k, a_i^k, y_i^k \right]
$$

$$
= \sum_{\tilde{a}_i^k, a_i^k, y_i^k} \left( \sum_{l=1}^T \delta_l^{-1} g_i (a(l)) + S_i (\tilde{a}_i^T, \rho^T (\tilde{a}_i^T, a_i^T, y_i^T)) \right) \cdot \Pr (\tilde{a}_i^T, a_i^T, y_i^T | \lambda_i^T, \alpha^T, \tilde{a}_i^k, a_i^k, y_i^k)
$$

(14)

denote player $i$’s expected payoff conditional on $(\tilde{a}_i^k, a_i^k, y_i^k)$; let

$$
v_i^T (\alpha^T, \rho^T, \mu^T; S, \tilde{a}_i^k, a_i^k, y_i^k) = \sum_{\tilde{a}_i^k, a_i^k, y_i^k} U_i^T (\alpha^T, \rho^T, \mu^T; S, \tilde{a}_i^k, a_i^k, y_i^k) \mu_i^T (\tilde{a}_{i-1}^k, a_{i-1}^k, y_{i-1}^k | \tilde{a}_i^k, a_i^k, y_i^k)
$$

(15)

denote player $i$’s expected payoff conditional on $(\tilde{a}_i^k, a_i^k, y_i^k)$. Write $U_i^T (\alpha^T, \rho^T; S)$ for $U_i^T (\alpha^T, \rho^T, \mu^T; S, \tilde{a}_i^0, a_i^0, y_i^0)$, and $v_i^T (\alpha^T, \rho^T, \mu^T; S)$ for $v_i^T (\alpha^T, \rho^T, \mu^T; S, \tilde{a}_i^0, a_i^0, y_i^0)$.

Given the mediator’s strategy $\lambda^T$, a strategy profile $(\alpha^T, \rho^T)$ is a Nash equilibrium if, for all $i \in N$ and all $(\alpha_i^T, \rho_i^T) \in A_i^T \times \Sigma_i^T$,

$$
U_i^T (\alpha^T, \rho^T; S) \geq U_i^T (\alpha_i^T, \rho_i^T; \alpha_{-i}^T, \rho_{-i}^T; S);
$$

(14)

As usual, $y_0^0$ denotes the null history $\emptyset$ and $Y_0^0$ the set whose only element is $y_0^0$.  

24
an assessment \((\alpha^T, \rho^T, \mu^T)\) is a perfect Bayesian equilibrium if

- for all \(i \in \mathcal{N}, k \leq T\), and \((\tilde{a}_i^k, a_i^k, y_i^k) \in A_i^k \times A_i^k \times Y_i^k\) such that

\[
\Pr \left( y_i^k \mid \lambda^{T \eta}, \alpha^T, a_i^k, y_i^k \right) > 0,
\]

the belief \(\mu^T_i (\cdot \mid \tilde{a}_i^k, a_i^k, y_i^k)\) is derived from \((\lambda^{T \eta}, \alpha^T_{-i})\) using Bayes’ rule;

- for all \(i \in \mathcal{N}, k \leq T\), \((\tilde{a}_i^k, a_i^k, y_i^k) \in A_i^k \times A_i^k \times Y_i^k\), and \((\alpha_i^T, \rho_i^T) \in A_i^T \times \Sigma_i^T\),

\[
v_i^T (\alpha^T, \rho^T, \mu^T; S, \tilde{a}_i^k, a_i^k, y_i^k) \geq v_i^T (\alpha_i^T, \rho_i^T, \alpha^T_{-i}, \rho^T_{-i}; S, \tilde{a}_i^k, a_i^k, y_i^k) . \quad (16)
\]

We say \(S\) enforces \((\lambda^{T \eta}, \alpha^T, \rho^T)\) if \((\alpha^T, \rho^T)\) is a Nash equilibrium, given the mediator’s strategy \(\lambda^{T \eta}\). The total per-period efficiency loss of \(S\) conditional on \((\lambda^{T \eta}, \alpha^T, \rho^T)\) is

\[
WL(T, \delta, S) \equiv - \sum_{i=1}^{n} \frac{1 - \delta}{1 - \delta T} E_{\tilde{a}_i^T, a_i^T, y_i^T} [S_i (\tilde{a}_i^T, \rho (\tilde{a}_i^T, a_i^T, y_i^T)) \mid \lambda^{T \eta}, \alpha^T] .
\]

To prove Proposition ??, we must show that for any \(\epsilon > 0\) it is possible to find \(S\) that satisfies (13) and enforces a Nash equilibrium with efficiency loss less than \(|L^*(\eta)| + \epsilon\).

### B.1.2 A transfer scheme enforcing \(\eta\) in the \(T\)-period game

Fix any \(\epsilon > 0\). By assumption we can choose \(x : A \times Y \times \mathcal{N} \to \mathbb{R}\) that strictly enforces \(\eta\) with efficiency loss less than \(|L^*(\eta)| + \epsilon/2\). Fix an \(\omega^* \in P\) that maximizes the total expected values \(\sum_{i=1}^{n} E [x_i(a, y) \mid \eta, \omega]\). Let

\[
w_1(a, y) \equiv x_1(a, y) + \sum_{i=2}^{n} E_{a', y'} [x_i(a', y') \mid \eta, P(a, y)],
\]

\[
w_i(a, y) \equiv x_i(a, y) - E_{a', y'} [x_i(a', y') \mid \eta, P(a, y)] \quad \text{for each } i \neq 1.
\]

This ensures that the expected value \(E_{a', y'} [w_i(a', y') \mid \eta, \omega]\) is highest conditional on \(\omega^*\) for each player \(i\), i.e.,

\[
\omega^* \in \arg \max_{\omega} E_{a', y'} [w_i(a', y') \mid \eta, \omega] \quad \text{for each } i.
\]
By Lemma 2 we can choose, for each \( w_i : A \times Y \to \mathbb{R} \) and each player \( j \neq i \), a function \( z_{ij} \) that satisfies the conditions of the lemma. Define

\[
    z_{ii} \equiv \frac{1}{n - 1} \sum_{j \neq i} z_{ij}.
\]

Building on \( w_i \) and \( z_{ij} \), we can now introduce the transfer mechanism \( S^{**} = (S_1^{**}, \ldots, S_n^{**}) \) that enforces \( \eta \) in the \( T \)-period game. \( S^{**} \) consists of four components. The first is a delayed communication mechanism similar the one in Fong, Gossner, Hörner, and Sannikov (2011). It requires that each player \( i \) pay a penalty equal to the difference between a target and the realized discounted value of \( w_i \) at the end of the \( T \)-period game whenever the latter is below the former. Denote the discounted value of \( w_i \) in the \( T \)-period game given the reports \((\tilde{a}_T, \tilde{y}_T)\) by

\[
    \Pi_i(\tilde{a}_T, \tilde{y}_T) \equiv \sum_{k=1}^{T} \delta^{k-1} w_i(\tilde{a}(k), \tilde{y}(k)).
\]

For each player \( i \) and each \((\tilde{a}_T, \tilde{y}_T) \in A_T \times Y_T\), let

\[
    S_i^*(\tilde{a}_T, \tilde{y}_T) \equiv -\max \{ K_i - \Pi_i(\tilde{a}_T, \tilde{y}_T), 0 \}, \tag{17}
\]

where

\[
    K_i \equiv E_{a,y}[w_i(a,y)|\eta,\omega^*] \sum_{k=1}^{T} \delta^{k-1} + T^{2/3}. \tag{18}
\]

Under \( S_i^* \), player \( i \)'s transfer is increasing one-to-one in the discounted value of \( w_i \) up to a cap \( K_i \).

Although the target \( K_i \) is chosen so that ex ante it is extremely unlikely that \( \Pi_i \) will exceed the target, there is a nonzero probability that player \( i \) learns through his own private signals that \( \Pi_i \) is likely to be above target, at which point his incentive to cooperate in the remaining periods breaks down. The second and the third components of our mechanism deal with the learning problem.

The second component of our mechanism is constructed so that whenever player \( i \) believes \( \Pi_i \) is likely to be above target, he also believes that he will almost surely receive a bonus that provides extra incentives from some player \( j \). However, we cannot directly make player \( j \) pay player \( i \) whenever \( \Pi_i \) is larger than
because doing so would merely transfer player $i$’s learning problem to player $j$—player $j$ would now want to deviate when he believes that $\Pi_i$ is large. We avoid this problem by making the additional reward conditional on the proxy variables $z_{ij}$. To this end, for each $i, j \in \mathcal{N}$ and $(a^T, y^T) \in A^T \times Y^T$, let

$$
\Gamma_{ij} (a^T, y^T) \equiv \sum_{k=1}^{T} \delta^{k-1} z_{ij} (a(k), y(k))
$$
denote the total discounted value of $z_{ij}$ in the $T$-period game. First, we raise the transfer of player $i$ by $\max \{\Pi_i (a^T, y^T) - K_i, 0\}$ when $\Gamma_{ii}$ is greater than $(K_i - \frac{1}{2}T^{2/3})$. To keep the total transfer negative, we also subtract the transfer of each player $j \neq i$ by $(\bar{\Pi}_j - K_j)$ when $\Gamma_{ij}$ is greater than $(K_i - \frac{1}{2}T^{2/3})$. By the definition of $z_{ii}$, when $\Gamma_{ii}$ is greater than $(K_i - \frac{1}{2}T^{2/3})$, there exists at least one player $j \neq i$ such that $\Gamma_{ij}$ is also greater than $(K_i - \frac{1}{2}T^{2/3})$. While we say player $j$ “pays” player $i$ a reward, what player $j$ actually pays is different from what player $i$ receives. To ensure that player $j$ has no incentive to misreport, the amount deducted from player $j$’s payoff is set independent of $\hat{y}_j^T$.

Hence, the second component of our mechanism for player $i$ is equal to

$$
L_i (\bar{a}^T, \hat{y}^T) \equiv \max \{\Pi_i (\bar{a}^T, \hat{y}^T) - K_i, 0\} f_{ii} (\bar{a}^T, \hat{y}^T) - \sum_{j \neq i} (\bar{\Pi}_j - K_j) f_{ji} (\bar{a}^T, \hat{y}^T),
$$

where

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

The first term on the right-hand side of (19) is the extra reward player $i$ may receive from the other players, while the second summation term is the total reward player $i$ may have to pay the other players.

Because $L_i$ depends on $\hat{y}^T$, player $i$ may try to manipulate $L_i$ by lying about his signals. The third component of the mechanism is used to induce player $i$ to report truthfully. For all $(\bar{a}^T, \hat{y}^T) \in A^T \times Y^T$, set

$$
D_i (\bar{a}^T, \hat{y}^T) \equiv T^{-2} \sum_{k=1}^{T} d_i (\bar{a}(k), \hat{y}(k)),
$$
Lemma 3. For any \((a, y) \in A \times Y\),
\[
d_i(a, y) \equiv 2p_i(y_{-i} \mid a, y_i) - \sum_{y'_{-i} \in Y_{-i}} (p_i(y'_{-i} \mid a, y_i))^2 - 1.
\]

The last component of our mechanism is a set of budget-balanced transfers. If \(\hat{y}(k) \notin Y^*\) for some \(k \leq T\), let
\[
F_i(\tilde{a}^T, \hat{y}^T) = -d_0 \quad \text{for each } i,
\]
for some \(d_0 > 0\) large enough. Otherwise, let
\[
F_i(\tilde{a}^T, \hat{y}^T) = -\sum_{i=2}^n \sum_{k=1}^T \delta^{k-1} E_{a,y} [x_i(a, y) \mid \eta, P(\tilde{a}(k), \hat{y}(k))], \tag{20}
\]
\[
F_i(\tilde{a}^T, \hat{y}^T) = \sum_{k=1}^T \delta^{k-1} E_{a,y} [x_i(a, y) \mid \eta, P(\tilde{a}(k), \hat{y}(k))] \text{ for each } i \neq 1. \tag{21}
\]

We can now define player \(i\)'s transfer as the combination of these four mechanisms,
\[
S_i^* (\tilde{a}^T, \hat{y}^T) = S_i^* (\tilde{a}^T, \hat{y}^T) + L_i (\tilde{a}^T, \hat{y}^T) + D_i (\tilde{a}^T, \hat{y}^T) + F_i (\tilde{a}^T, \hat{y}^T). \tag{22}
\]
Note that by construction for any \((\tilde{a}^T, \hat{y}^T) \in A^T \times Y^T\),
\[
\sum_{i=1}^n F_i (\tilde{a}^T, \hat{y}^T) \leq 0, \quad \sum_{i=1}^n (S_i^* (\tilde{a}^T, \hat{y}^T) + L_i (\tilde{a}^T, \hat{y}^T)) \leq 0 \quad \text{and} \quad \sum_{i=1}^n D_i (\tilde{a}^T, \hat{y}^T) \leq 0.
\]
Hence, the total transfer is always negative.

Denote the action strategy that chooses the recommended action \(\tilde{a}_i(k)\) in every period \(k\) by \(\alpha_i^{T^*}\), the truth-telling reporting strategy that always reports \(y_i^T\) by \(\rho_i^{T^*}\), and the strategy profile where every player \(i\) chooses \((\alpha_i^{T^*}, \rho_i^{T^*})\) by \((\alpha^{T^*}, \rho^{T^*})\). Let \(\lambda^{T^}\) denote the strategy for the mediator of always recommending according to \(\eta\).

We claim that when \(T\) is large and \(\delta\) is close to 1, the transfer functions \(S^{**}\) enforces \((\lambda^{T^}, \alpha^{T^*}, \rho^{T^*})\) with efficiency loss \(WL(T, \delta, S^{**})\) close to \(L^*(\eta)\).

Lemma 3. For any \(\epsilon > 0\), there exists a \(T_0\) such that, for all \(T \geq T_0\) and all \(\delta \geq 1 - T^{-2}\),
1. \(WL(T, \delta, S^{**}) \leq |L^*(\eta)| + \epsilon\), and
2. \((\alpha^{T^*}, \rho^{T^*})\) is a Nash equilibrium of \(\Gamma^T(G_0, S^{**}, \delta)\), given \(\lambda^{T^}\).
**B.1.3 Proof of Lemma 3**

In each period $k$ of the $T$-period game, conditional on any $(\tilde{a}_i(k), a_i(k), y_i(k))$, player $i$ believes that the other players have received the recommended actions in the set

$$\tilde{A}_{-i}(\tilde{a}_i(k), a_i(k), y_i(k)) \equiv \{ a_{-i} \in A_{-i} : \eta(a_{-i}, \tilde{a}_i(k)) > 0, p_i(y_i(k)|a_{-i}, a_i(k)) > 0 \}.$$ 

If $\tilde{A}_{-i}(\tilde{a}_i(k), a_i(k), y_i(k))$ is empty, then player $i$ knows that some player must have deviated from the recommended actions.

To prove the lemma, it suffices to show that conditional on any history $(\tilde{a}_i^k, a_i^k, y_i^k)$ such that $a_i(l) = \tilde{a}_i(l)$ and $\tilde{A}_{-i}(\tilde{a}_i(l), a_i(l), y_i(l))$ is nonempty for each $l \leq k$, it is optimal for to choose the recommended action in period $(k + 1)$. Conditional on $(\tilde{a}_i^k, a_i^k, y_i^k)$, player $i$'s belief over the other players' private information $(\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k)$ is given by

$$\Pr(\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k|\lambda^{T\eta}, \alpha^T_{-i}, \tilde{a}_i^k, a_i^k, y_i^k).$$

Hence, it suffices to show that for all $(\alpha_i^T, \rho_i^T) \in A_i^T \times \Sigma_i^T$,

$$\sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T(\alpha_i^T, \rho_i^T; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr(\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k|\lambda^{T\eta}, \alpha^T_{-i}, \tilde{a}_i^k, a_i^k, y_i^k)$$

$$\geq \sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T(\alpha_i^T, \rho_i^T; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr(\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k|\lambda^{T\eta}, \alpha^T_{-i}, \tilde{a}_i^k, a_i^k, y_i^k).$$

(23)

By the definition of $S^{**}$, it suffices to consider $\rho_i^T$ such that $\rho_i^T(\tilde{a}_i^T, a_i^T, y_i^T) \in Y_i^T$ for all $(\tilde{a}_i^T, a_i^T, y_i^T)$.

Substitute (22) into (15) and rearrange terms. We have

$$\sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T(\alpha_i^T, \rho_i^T; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr(\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k|\lambda^{T\eta}, \alpha^T_{-i}, \tilde{a}_i^k, a_i^k, y_i^k)$$

$$= V_i(\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) - \sum_{l=1}^{3} R_i(\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) - K_i,$$

(24)

29
where

\[
V_i (\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) \equiv E_{a^T, \omega_T} \left[ \sum_{l=1}^{T} \delta^{l-1} g_l (a(l)) + \Pi_i (\tilde{a}_i^T, \rho_i^T (\tilde{a}_i^T, a_i^T, y_i^T), y_i^T) + F_i (\omega_T^T) \right] \lambda_T^i, \alpha_i^T
\]

\[
R_i^1 (\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) \equiv E_{a^T, \omega_T} \left[ \max \left\{ \Pi_i (\tilde{a}_i^T, \rho_i^T (\tilde{a}_i^T, a_i^T, y_i^T), y_i^T) - K_i, 0 \right\} (1 - f_i (\tilde{a}_i^T, \rho_i^T (\tilde{a}_i^T, y_i^T)) \right] \lambda_T^i, \alpha_i^T
\]

\[
R_i^2 (\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) \equiv \sum_{j \neq i} E_{a^T, \omega_T} \left[ (\Pi_j - K_j) f_{ji} (\tilde{a}_i^T, \rho_i^T (\tilde{a}_i^T, a_i^T, y_i^T), y_j^T, y_i^T) \right] \lambda_T^i, \alpha_i^T, \alpha_j^T, \tilde{a}_i^k, a_i^k, y_i^k
\]

\[
R_i^3 (\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) \equiv -E_{a^T, \omega_T} \left[ D_i (\tilde{a}_i^T, \rho_i (\tilde{a}_i^T, a_i^T, y_i^T), y_i^T) \right] \left| \lambda_T^i, \alpha_i^T, \alpha_j^T, \tilde{a}_i^k, a_i^k, y_i^k \right|
\]

The term \( V_i \) in (24) is the sum of player \( i \)'s own expected stage-game payoff and the expected value of the transfer \( x_i \) in the \( T \)-period game. By assumption, this term is maximized when player \( i \) follows the equilibrium strategy \( \alpha_i^T \). Let

\[
\Delta \equiv \min \left\{ \sum_{a_{-i}} (g_i (\tilde{a}) + E_y [w_i (\tilde{a}, y)|\tilde{a}]) \eta (\tilde{a}) - \sum_{a_{-i}} (g_i (a_i', \tilde{a}_i) + E_y [w_i (\tilde{a}, \rho_i (y_i), y_{-i})|a_i', \tilde{a}_i]) \eta (\tilde{a}) \mid \begin{array}{c}
i \in N, \tilde{a}_i \in \text{supp}(\eta_i), a_i' \in A_i, \rho_i' \in \Psi_i \end{array} \right\}
\]

denote the minimum loss in the sum of player \( i \)'s expected stage-game payoff and his expected reward when he deviates unilaterally. Let \( A_i^T (\tilde{a}_i^k, a_i^k, y_i^k) \) denote the set of player \( i \)'s action strategies that prescribe \( a_i \neq \tilde{a}_i(k+1) \) in period \( (k+1) \) for some \( \tilde{a}_i(k+1) \), conditional on \( (\tilde{a}_i^k, a_i^k, y_i^k) \). Then

\[
V_i (\alpha_i^{T^*}, \rho_i^{T^*}; \tilde{a}_i^k, a_i^k, y_i^k) - V_i (\alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k) \geq \delta^k \Delta.
\]

The term \( R_i^1 \) in (24) denotes the expected net truncation effect. Recall that \( S_i^k \) increases linearly in the reward \( \Pi_i (\tilde{a}_i^T, \tilde{y}_i^T) \) only up to \( K_i \). The first component of \( L_i \) pays player \( i \) an additional reward equal to \( \max \{ \Pi_i (\tilde{a}_i^T, \tilde{y}_i^T) - K_i, 0 \} \), the truncated amount, when \( f_{ii} (\tilde{a}_i^T, \tilde{y}_i^T) = 1 \). Hence, the cap \( K_i \) will be binding for player \( i \) only when \( \Pi_i (\tilde{a}_i^T, \tilde{y}_i^T) > K_i \) and \( f_{ii} (\tilde{a}_i^T, \tilde{y}_i^T) = 0 \). When the players report truthfully, this event occurs when

\[
\Gamma_i (\tilde{a}_i^T, \tilde{y}_i^T) + \frac{1}{2} T^{2/3} < K_i < \Pi_i (\tilde{a}_i^T, \tilde{y}_i^T).
\]

(26)
By condition 1 of Lemma 1,

\[ E_{\alpha^T, a^T, y^T} [\Gamma_{ii} (\tilde{a}^T, y^T) - \Pi_i (\tilde{a}^T, y^T)] = 0. \]

Let

\[ c_2 = \max_{i,a,y} |w_i(a, y) - z_{ii}(a, y)|. \]

Substituting \( \delta^{-1} (w_i(\tilde{a}(l), y(l)) - z_{ii}(\tilde{a}(l), y(l))) \) for \( x_1, \frac{1}{2}T^{2/3} \) for \( d, T \) for \( l_0 \), and \( c_2 \) for \( \nu \) in Hoeffding’s inequality we have

\[
\Pr \left( \left\{ (\tilde{a}^T, a^T, y^T) : \Pi_i (\tilde{a}^T, y^T) - \Gamma_{ii} (\tilde{a}^T, y^T) > \frac{1}{2}T^{2/3} \right\} \right| \lambda^T, \alpha^T, \tilde{a}_i^k, a_i^k, y_i^k \leq \exp \left( -\frac{1}{8c_2^2}T^{1/3} \right). \]

Since \( \max \{ \Pi_i (\tilde{a}^T, y^T) - K_i, 0 \} \leq c_1T \) and (26) holds only when \( \Pi_i (\tilde{a}^T, y^T) - \Gamma_{ii} (\tilde{a}^T, y^T) > \frac{1}{2}T^{2/3} \),

\[ R_i^1 (\alpha_i^T, \rho_i^T, \tilde{a}_i^k, a_i^k, y_i^k) \leq c_1T \exp \left( -\frac{1}{8c_2^2}T^{1/3} \right). \quad (27) \]

The second component of \( L_i \) deducts \( (\tilde{P}_j - K_j) \) from player \( i \)’s payoff whenever \( \Gamma_{ji} \) is greater than \( (K_j - \frac{1}{2}T^{2/3}) \) for some \( j \neq i \). The term \( R_i^2 \) in (24) denotes the expected amount deducted from player \( i \)’s through this channel. By condition 2 of Lemma 1, and the definition of \( K_j \),

\[
\sum_{l=1}^{k} \delta^{-1} E_{\tilde{a}(l), y(l)} [w_j (\tilde{a}(l), y(l)) | \eta, \omega(l)] + \sum_{l=k+1}^{T} \delta^{-1} E_{\tilde{a}(l), y(l)} [w_j (\tilde{a}(l), y(l)) | \eta] \leq K_j - T^{2/3}. \]

Let \( c_3 = \max_{i,j,y} |z_{ij}(\tilde{a}(l), y(l))| \). Substituting \( \delta^{-1} z_{ji}(\tilde{a}(l), y(l)) \) for \( x_1, T^{2/3} \) for \( d, T \) for \( l_0 \), and \( c_3 \) for \( \nu \) in Hoeffding’s inequality, we have

\[
\Pr \left( \left\{ (\tilde{a}^T, a^T, y^T) : \Gamma_{ji} (\tilde{a}^T, y^T) > K_j \right\} \right| \lambda^T, \alpha^T, \tilde{a}_i^k, a_i^k, y_i^k \leq \exp \left( -\frac{1}{2c_3^2}T^{1/3} \right). \]

It follows that

\[ R_i^2 (\alpha_i^T, \rho_i^T, \tilde{a}_i^k, a_i^k, y_i^k) \leq c_1T \exp \left( -\frac{1}{2c_3^2}T^{1/3} \right). \quad (28) \]
To bound $D_i$, let

$$c_4 = \max_{i,a,y} \left| 2p_{-i}(y_{-i} | a, y_i) - \sum_{y'_{-i} \in Y_{-i}} (p_{-i}(y'_{-i} | a, y_i))^2 - 1 \right|.$$ 

Hence

$$D_i \left( \tilde{a}^T, y^T \right) \geq T^{-2} \cdot T \cdot (-c_4) = -c_4 T^{-1}. \quad (29)$$

It follows that

$$R_i^3 \left( \alpha_i^{T*}, \rho_i^{T*}, \tilde{a}_i^k, a_i^k, y_i^k \right) \leq c_4 T^{-1}. \quad (30)$$

Combining all these bounds, and noting that $R_i^l \left( \alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k \right) \geq 0$ for each $l$, we have

$$\sum_{l=1}^{3} R_i^l \left( \alpha_i^{T*}, \rho_i^{T*}; \tilde{a}_i^k, a_i^k, y_i^k \right) - \sum_{l=1}^{3} R_i^l \left( \alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k \right) \leq c_1 T \exp \left( -\frac{1}{8c_2^2} T^{1/3} \right) + c_1 T \exp \left( -\frac{1}{2c_3^2} T^{1/3} \right) + c_4 T^{-1}. \quad (31)$$

Set $\delta^*(T) \equiv 1 - T^{-2}$. This ensures that $\left( \delta^*(T) \right)^T$ tends to 1 as $T$ tends to infinity. Note that the right-hand side of (31) tends to zero as $T$ tends to infinity. So we can choose $T_1$ large enough such that for all $T \geq T_1$ and $\delta \geq \delta^*(T)$,

$$\delta^T \Delta \geq \left( \delta^*(T) \right)^T \Delta \geq c_1 T \exp \left( -\frac{1}{8c_2^2} T^{1/3} \right) + c_1 T \exp \left( -\frac{1}{2c_3^2} T^{1/3} \right) + c_4 T^{-1}.$$ 

It follows that

$$V_i \left( \alpha_i^{T*}, \tilde{a}_i^k, a_i^k, y_i^k \right) \equiv V_i \left( \alpha_i^T, \tilde{a}_i^k, a_i^k, y_i^k \right) \geq \sum_{l=1}^{3} R_i^l \left( \alpha_i^{T*}, \rho_i^{T*}; \tilde{a}_i^k, a_i^k, y_i^k \right) - \sum_{l=1}^{3} R_i^l \left( \alpha_i^T, \rho_i^T; \tilde{a}_i^k, a_i^k, y_i^k \right).$$

Thus, for $\left( \alpha_i^T, \rho_i^T \right) \in A_i^T \left( \tilde{a}_i^k, a_i^k, y_i^k \right) \times \Sigma_i^T$, $\left( \tilde{a}_i^k, a_i^k, y_i^k \right) \in A_i^k \times A_i^k \times Y_i^k$ with $a_i(l) = \tilde{a}_i(l)$ and $\overline{A}_{-i} \left( \tilde{a}_i(l), a_i(l), y_i(l) \right)$ is nonempty for each $l \leq k$, and $k = 0, 1,$
\[ \ldots, (T - 1), \]
\[ \sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T (\alpha_{T^*}, \rho_{T^*}; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr (\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k | \lambda^{T^*}, \alpha_T^{T^*}, \tilde{a}_i^k, a_i^k, y_i^k) \]
\[ \geq \sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T (\alpha_{T^*}, \rho_T^*; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr (\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k | \lambda^{T^*}, \alpha_T^{T^*}, \tilde{a}_i^k, a_i^k, y_i^k). \]

This means that assuming that other players choose \((\alpha_{T^*}, \rho_{T^*})\), it is never a best response for player \(i\) to deviate from \(\alpha_i^{T^*}\).

To prove that \((\alpha_{T^*}, \rho_{T^*})\) maximizes \(U_i^T (\alpha_T^*, \rho_T^*; S^{**})\) with respect to \((\alpha_T^*, \rho_T^*)\), we still need to show that, for all \(\rho_T^* \in \Sigma_T\),
\[ \sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T (\alpha_{T^*}, \rho_T^*; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr (\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k | \lambda^{T^*}, \alpha_T^{T^*}, \tilde{a}_i^k, a_i^k, y_i^k) \]
\[ \geq \sum_{\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k} U_i^T (\alpha_T^*, \rho_T^*; S^{**}, \tilde{a}_i^k, a_i^k, y_i^k) \Pr (\tilde{a}_{-i}^k, a_{-i}^k, y_{-i}^k | \lambda^{T^*}, \alpha_T^{T^*}, \tilde{a}_i^k, a_i^k, y_i^k). \]

When player \(i\) follows \(\alpha_i^{T^*}\), the reporting strategy \(\rho_i^T\) affects only \(R_l^i, l = 1, 2, 3\), but not \(V_i\). Write \(\tilde{a}_i^T\) for \(\rho_i^T (\tilde{a}_i^T, a_i^T, y_i^T)\). There are two cases to consider.

Case 1. Suppose \(p_{-i}(\cdot | \tilde{a}(k), \tilde{y}_i(k)) = p_{-i}(\cdot | \tilde{a}(k), y_i(k))\) for each \(\tilde{a}_{-i}(k)\) such that \(\eta(\tilde{a}_{-i}(k) | \tilde{a}_i(k)) > 0\) and each \(k \in \{1, \ldots, T\}\). In this case, for \(l = 1, 2, 3\),
\[ R_l^i (\alpha_i^{T^*}, \rho_i^T; \tilde{a}_i^T, a_i^T, y_i^T) = R_l^i (\alpha_i^{T^*}, \rho_i^T; \tilde{a}_i^T, a_i^T, y_i^T). \]

The case for \(l = 1, 2\) follows from conditions 1 and 2 of Lemma 2. The case for \(l = 3\) follows from the definition of \(D_i\).

Case 2. Suppose \(p_{-i}(\cdot | \tilde{a}(k), \tilde{y}_i(k)) \neq p_{-i}(\cdot | \tilde{a}(k), y_i(k))\) for some \(k \in \{1, \ldots, T\}\) and \(\tilde{a}_{-i}(k)\) such that \(\eta_{-i}(\tilde{a}_{-i}(k) | \tilde{a}_i(k)) > 0\). It is a standard result in the scoring-rule literature that
\[ E_{y_{-i}(k)} [d_i(\tilde{a}(k), \tilde{y}_i(k)) | \tilde{a}(k), y_i(k)] < E_{y_{-i}(k)} [d_i(\tilde{a}(k), y_i(k), y_{-i}(k)) | \tilde{a}(k), y_i(k)]. \]

Hence,
\[ R_l^i (\alpha_i^{T^*}, \rho_i^T; \tilde{a}_i^T, a_i^T, y_i^T) - R_l^i (\alpha_i^{T^*}, \rho_i^T; \tilde{a}_i^T, a_i^T, y_i^T) \geq T^{-2} \cdot \min \{ \eta_{-i}(\tilde{a}_{-i} | \tilde{a}_i) > 0 | \tilde{a} \}, \]

33
where
\[ \bar{d} \equiv \min \{ E_{y_0} [d_i(a, y_1, y_2)] | a, y_1 | - E_{y_0} [d_i(a, \tilde{y}_i, y_2)] | a, y_1 | \} \]
\[ a \in A, y_1, \tilde{y}_1 \in Y \text{ and } p_{-1}(\cdot | a, \tilde{y}_1) \neq p_{-1}(\cdot | a, y) \} > 0. \]

We have already shown that
\[ \sum_{l=1}^{2} R_i^l \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right) - \sum_{l=1}^{2} R_i^l \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right) \]
\[ \leq c_1 T \exp \left( -\frac{1}{8c_2^2} T^{1/3} \right) + c_1 T \exp \left( -\frac{1}{2c_3^2} T^{1/3} \right). \quad (34) \]

Since the right-hand side of (34) decays faster than \( T^{-2} \), we can choose \( T_2 \) large enough such that for all \( T \geq T_2 \) and all \( \delta \),
\[ R_i^3 \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right) - R_i^3 \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right) \]
\[ \geq \sum_{l=1}^{2} R_i^l \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right) - \sum_{l=1}^{2} R_i^l \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right). \]

We now turn to the first part of the lemma. By definition we have
\[ E_{a_*^T, a_*^T, y_T} \left[ S_{\alpha_*^T} \left( a_*^T, y_T \right) \right] \]
\[ = E_{a_*^T, a_*^T, y_T} \left[ \Pi_i \left( a_*^T, y_T \right) + F_i \left( a_*^T, y_T \right) - K_i \left( a_*^T, y_T \right) \right] - \sum_{i=1}^{3} R_i^l \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right) \]
\[ = 1 - \frac{1 - \delta^T}{1 - \delta} \left( E_{a_*^T, y_T} [x_i (a_*^T, y_T) | \eta, \omega^*] - E_{a_*^T, y_T} [x_i (\tilde{a}, y_T) | \eta] \right) \]
\[ - T^{2/3} - \sum_{l=1}^{3} R_i^l \left( \alpha_{i_*}^T, \rho_{i_*}^T; a_{i_*}^T, y_i^T \right). \]

Since \( (\delta^*(T))^T \) tends to 1 as \( T \) tends to infinity, \( (1 - \delta^*(T)) \left( 1 - (\delta^*(T))^T \right)^{-1} \)
is of the order of \( T^{-1} \). It follows that
\[ \frac{1 - \delta^*(T)}{1 - (\delta^*(T))^T} T^{2/3} \]
tends to 0 as \( T \) tends to infinity. By (27), (28) and (30) each \( R_i^l \) is bounded by a term that tends to zero as \( T \) tends to infinity. Hence, for any \( \epsilon \) we can choose \( T_3 \).
large enough such that for all \( T \geq T_3 \) and \( \delta \geq \delta^*(T) \), and for each \( i \),

\[
\frac{1 - \delta}{1 - \delta T} E_{\tilde{a}, y} \left[ S_i^{**} \left( \tilde{a}^T, y^T \right) | \lambda^T \eta, \alpha^T \right] - E_{\tilde{a}, y} \left[ x_i \left( \tilde{a}, y \right) | \eta \right] + \frac{\epsilon}{2n}.
\]

(35)

Set \( T_0 = \max\{T_1, T_2, T_3\} \). Then (32) and (33) hold for all \( T \geq T_0 \) and \( \delta \geq \delta^*(T) \), and by (35),

\[
WL(T, \delta, S^{**} \leq \sum_{i=1}^{n} \left( E_{\tilde{a}, y} \left[ x_i \left( \tilde{a}, y \right) | \eta, \omega^* \right] - E_{\tilde{a}, y} \left[ x_i \left( \tilde{a}, y \right) | \eta \right] \right) + \frac{\epsilon}{2} \leq L^*(\eta) + \epsilon.
\]

B.2 Step 2: Implementation

B.2.1 Equilibrium structure

We prove Theorem 1 by construction. The availability of side-payments allows us to construct an equilibrium that involves only two states: a cooperative state and a noncooperative state. The players start off in the cooperative state, which lasts for \( T \) periods.

In the cooperative state, the mediator sends the recommendations according to the distribution \( \eta \) in each period, does not reveal any further information in the first \( (T - 1) \) periods, and reveal all the recommendations publicly at the end of period \( T \). The players follow the following strategy profile. On the equilibrium path, they follow the recommendations for \( T \) periods; in the first \( (T - 1) \) periods, they announce constant messages, and do not exchange side-payments; at the end of period \( T \), they report their private signals in all \( T \) periods truthfully. Off the equilibrium path—if a player has deviated from the recommended action to him or he has observed deviations by the other players, then in the remaining periods he follows a different strategy to be specified in detail later.

In the cooperative state, the mediator sends the recommendations according to the distribution \( \eta \) in each period, does not reveal any further information in the first \( (T - 1) \) periods, and reveal all the recommendations publicly at the end of period \( T \). The players follow the following strategy profile. On the equilibrium path, they follow the recommendations for \( T \) periods; in the first \( (T - 1) \) periods, they announce constant messages, and do not exchange side-payments; at the end of period \( T \), they report their private signals in all \( T \) periods truthfully. Off the equilibrium path—if a player has deviated from the recommended action to him or he has observed deviations by the other players, then in the remaining periods he follows a different strategy to be specified in detail later.

On the basis of recommended actions and the reports of the players at the end of period \( T \), the players exchange side-payments and determine the probability of switching to the noncooperative state in the next period using the public randomization device. If the players stay in the cooperative state, they repeat the same process. If they switch to the noncooperative state, which is absorbing, they play the stage-game Nash equilibrium \( a^N \) forever. In the noncooperative state,
the mediator sends constant recommendation and announces constant message; the players ignore the recommendations, send constant messages and do not make side-payments.

To complete the description of the equilibrium strategy profile, we still need to specify the strategies of the players off the equilibrium path, and the side-payments and the transition probabilities after $T$ periods in the cooperative state.

**B.2.2 Enforcing $\eta$ in the $T$-period game by a perfect Bayesian equilibrium**

To describe the off-the-equilibrium-path strategies in the perfect Bayesian equilibrium of the original infinitely-repeated game, we now show that $\eta$ can be enforced in the $T$-period game by a perfect Bayesian equilibrium by slightly modifying the transfer functions $S_{**}$ and strategy profile $(\alpha^T_*, \rho^T_*)$.

Choose $d_1 \geq 0$ large enough such that, for all $(\tilde{a}^T, \hat{y}^T)$,

$$-d_1 + \Pi_i (\tilde{a}^T, \hat{y}^T) \leq 0; \quad (36)$$

$$-d_1 + \Pi_i (\tilde{a}^T, \hat{y}^T) \geq S_{**}^i (\tilde{a}^T, \hat{y}^T). \quad (37)$$

Define a new transfer function for each player $i$ by setting, for each $(\tilde{a}^T, \hat{a}^T, \hat{y}^T) \in A^T \times A^T \times Y^T$,

$$S_i^\dagger (\tilde{a}^T, \hat{a}^T, \hat{y}^T) \equiv \begin{cases} S_{**}^i (\tilde{a}^T, \hat{y}^T), & \text{if } \tilde{a}^T = \hat{a}^T \text{ and } p (\hat{y}(k)|\tilde{a}(k)) > 0 \text{ for each } k; \\ -d_1 + \Pi_i (\tilde{a}^T, \hat{y}^T), & \text{if } \tilde{a}^T \neq \hat{a}^T; \\ -\delta^T v_i - g_i (a^N) \frac{1 - \delta}{1 - \delta}, & \text{if } \tilde{a}^T = \hat{a}^T \text{ and } p (\hat{y}(k)|\tilde{a}(k)) = 0 \text{ for some } k. \end{cases} \quad (38)$$

For each $m \notin A^T \times A^T \times Y^T$,

$$S_i^\dagger (m) = -\delta^T v_i^\ast - g_i (a^N) \frac{1 - \delta}{1 - \delta}.$$
\( \tilde{a}^T = \tilde{a}^T \) and no deviation is detected \( (p(\tilde{y}(k)|\tilde{a}(k)) > 0 \) for each \( k \), then the payoff to each player \( i \) is given by the original transfer function \( S^*_i \). If any player reports a deviation from the recommendations \( (\tilde{a}^T \neq \tilde{a}^T) \), then the payoff function of each player \( i \) is linear in the discounted value of \( w \) in the \( T \)-period game. If a deviation is detected \( (p(\tilde{y}(k)|\tilde{a}(k)) = 0 \) for some \( k \) but no one has admitted a deviation \( (\tilde{a}^T = \tilde{a}^T) \), then all players are punished. Here \( v^*_i \) is the average equilibrium payoff of the players in the infinitely-repeated game and the punishment will be implemented by having the players switch to the noncooperative state in the next period.

For each \( k \leq T \), let \( D_{i,k} \) denote the set of private histories \( (\tilde{a}^T_i, a^T_i, y^T_i) \) for player \( i \) in which he has not observed any deviations by the other players and, if he has deviated, he believes that there is a way to misreport his private signals such that the deviations are not detectable by the other players. That is, for each \( l \leq k \), \( \overline{A}_i(l) = (\tilde{a}_i(l), a_i(l), y_i(l)) \) is nonempty and, if \( a_i(l) \neq \tilde{a}_i(l) \), then there exists \( \tilde{y}_i(l) \) such that for all \( \tilde{a}_{-i}(l) \in \overline{A}_{-i}(\tilde{a}_i(l), a_i(l), y_i(l)) \),

\[
\text{supp } p_{-i}(\cdot|\tilde{a}_{-i}(l), a_i(l), y_i(l)) \subseteq \text{supp } p_{-i}(\cdot|\tilde{a}_{-i}(l), \tilde{a}_i(l), \tilde{y}_i(l)).
\]

The left-hand side of (39) is the set of signals that player \( i \) believes the other players have observed, the right-hand side is the set of signals that the other players believe is possible, assuming player \( i \)’s report is truthful. If (39) holds, then player \( i \) believes his deviation will not be detected.

Define a new reporting strategy \( \rho_i^{T+i} \) for each player \( i \) by setting, for each \( (\tilde{a}^T_i, a^T_i, y^T_i) \in A^T_i \times A^T_i \times Y^T_i \),

\[
\rho_i^{T+i} (\tilde{a}^T_i, a^T_i, y^T_i) \equiv \begin{cases} 
(\tilde{a}^T_i, \rho_i^{T+i} (\tilde{a}^T_i, a^T_i, y^T_i)) , & \text{if } (\tilde{a}^T_i, a^T_i, y^T_i) \in D_{i,T}; \\
(a^T_i, y^T_i), & \text{otherwise.}
\end{cases}
\] (40)

where \( \rho_i^{T+i} (\tilde{a}^T_i, a^T_i, y^T_i) \in Y^T_i \) is an optimal report of private signals under \( S^{**} \), assuming that the other players have followed \( (\alpha^{**}_i, \rho^{**}_i) \), i.e.,

\[
\rho_i^{T+i} (\tilde{a}^T_i, a^T_i, y^T_i) \\
\in \text{arg max } \sum_{\tilde{a}^T_{-i}, a^T_{-i}, y^T_{-i}} S^{**} (\tilde{a}^T, \tilde{y}^T_i, y^T_{-i}) \Pr (\tilde{a}^T_{-i}, a^T_{-i}, y^T_{-i} | \tilde{a}^T, a^T_i, y^T_i, \lambda) \cdot \alpha^{**}_i (\tilde{a}^T_i, a^T_i, y^T_i).}
\]

37
Define a new action strategy $\alpha_i^{T\dagger}$ for each player $i$ by backward induction. Suppose that $k = T$, or $k < T$ and $\alpha_i^{T\dagger} (\tilde{a}_i^{k+1}, a_i^l, y_i^l)$ has been defined for each $(\tilde{a}_i^{k+1}, a_i^l, y_i^l)$ and each $l \geq k$. For each $(\tilde{a}_i^{k-1}, a_i^{k-1}, y_i^{k-1})$, if $\tilde{a}_i^{k-1}, a_i^{k-1}, y_i^{k-1} \not\in D_{i,k-1}$, we let

$$
\alpha_i^{T\dagger} (\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1}) = \tilde{a}_i(k) \quad \text{for all } \tilde{a}_i(k) \in A_i.
$$

Otherwise, $\tilde{a}_i^{k-1}, a_i^{k-1}, y_i^{k-1} \in D_{i,k-1}$. Then for all $(a_i(k), y_i(k))$ and $\tilde{a}(k)$ such that $p_i(y_i(k) | \tilde{a}_i(k), a_i(k)) \eta(\tilde{a}(k)) > 0$, we have $\tilde{a}_i^k, a_i^k, y_i^k \in D_{i,k}$. Hence the probability

$$
\Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda, \alpha^T, \alpha_{-i}^T, \alpha_{-i}^{T\dagger}, a_i^k, y_i^k \right)
$$

is well-defined for all $(\tilde{a}^T, a^T, y^T)$. Conditional on $(\tilde{a}_i^k, a_i^k, y_i^k)$, if player $i$ follows $\alpha_i^{T\dagger}$ and $\rho_i^{T\dagger}$ and other players follow $(\alpha_{-i}^T, \rho_{-i}^{T\dagger})$, player $i$’s expected payoff is

$$
\pi_i^T \left( (\alpha_i^{T\dagger}, \rho_i^T, a_i^{T\dagger}, \alpha_{-i}^T, \rho_{-i}^{T\dagger}, S^T, \tilde{a}_i^k, a_i^k, y_i^k) \right) =
$$

$$
\sum_{\tilde{a}^T, a^T, y^T} \left( \sum_{l=1}^{T} \delta^{l-1} g_i (a(l)) + S_i^T (\tilde{a}^T, a^{T\dagger}, a^T, y^T) \right)
$$

$$
\cdot \Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda, \alpha_i^{T\dagger}, \alpha_{-i}^T, \tilde{a}_i^k, a_i^k, y_i^k \right).
$$

(41)

we define $\alpha_i^{T\dagger} (\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1})$ to be a solution to

$$
\max_{a_i(k)} \sum_{a_{-i}(k), y_i(k)} \pi_i^T \left( (\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1}) \right)
$$

$$
\alpha_i^{T\dagger} (\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1}) = \tilde{a}_i(k).
$$

(42)

**Claim 1.** Suppose that $a_i^{k-1} = \tilde{a}_i^{k-1}$ and, for each $l \leq k - 1$, $A_i (\tilde{a}_i(l), a_i(l), y_i(l))$ is nonempty. That is, player $i$ has not deviated and has not observed any deviations conditional on $(\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1})$. Then we have

$$
\alpha_i^{T\dagger} (\tilde{a}_i^k, a_i^{k-1}, y_i^{k-1}) = \tilde{a}_i(k).
$$

**Proof.** We show by backward induction. Suppose that $k = T$, or $k < T$ and the claim is true for all $l > k$. Then, for all $(\tilde{a}^T, a^T, y^T) \in A^T \times A^T \times Y^T$,

$$
\Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda, \alpha_i^{T\dagger}, \alpha_{-i}^T, \tilde{a}_i^k, a_i^k, y_i^k \right) = \Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda, \alpha_i^{T\dagger}, \tilde{a}_i^k, a_i^k, y_i^k \right),
$$

38
and if \( \breve{a}^T \neq a^T \),

\[
\Pr (\breve{a}^T, a^T, y^T | \lambda^{T\eta}, \alpha^{T*}, \breve{a}^k, a^k_i, y^k_i) = 0.
\]

It follows that

\[
\overline{v}_i^T (\alpha_i^{T*}, \rho_i^{T*}, \alpha_i^{T*}, \rho_i^{T*}; S^i, \breve{a}^k_i, a^k_i, y_i^k) = \overline{v}_i^T (\alpha_i^{T*}, \rho_i^{T*}; S^{**}, \breve{a}^k_i, a^k_i, y_i^k).
\]

By Lemma 3 the optimal solution to (42) is to follow the recommendation. \( \square \)

We define a system of beliefs \( \mu^{T*} \) as follows. For all \( i \in \mathcal{N}, k \leq T \), and \( (\breve{a}^k_i, a^k_i, y^k_i) \in A^k_i \times A^k_i \times Y^k_i \), if

\[
\Pr (y^k_i | \lambda^{T\eta}, \alpha^{T*}, \breve{a}^k_i, a^k_i) > 0,
\]

the belief \( \mu^{T*}_i (\cdot | \breve{a}^k_i, a^k_i, y^k_i) \) is derived from \( (\lambda^{T\eta}, \alpha^{T*}) \) using Bayes’ rule; otherwise, let \( \mu_i^{T*} (\cdot | \breve{a}^k_i, a^k_i, y^k_i) \) be any fixed distribution on \( A^k_i \times A^k_i \times Y^k_i \). The beliefs of the equilibrium paths will be irrelevant in checking the optimality of \( (\alpha^T, \rho^T) \).

**Lemma 4.** For any \( \epsilon > 0 \), there exists a \( T_1 \) such that, for all \( T \geq T_1 \) and all \( \delta \geq 1 - T^{-2} \),

1. \( \alpha_i^{T*} (\breve{a}^k_i, a^k_i, y^k_i) = \tilde{a}_i(k) \) whenever \( a_i^{k-1} = \tilde{a}_i^{k-1} \),

2. \( WL (T, \delta, S^i) \leq |L^*(\eta)| + \epsilon \), and

3. \( (\alpha^T, \rho^T, \mu^T) \) is a perfect Bayesian equilibrium of \( \Gamma^T (G_0, S^i, \delta) \), given \( \lambda^{T\eta} \).

**Proof.** Part 1 follows from Claim 1 and the definition of \( \alpha_i^{T*} \). Part 2 follows from Part 1, the definition of \( S^i \), and Lemma 3. It remains to prove Part 3. We first show that the action strategy \( \alpha^{T*} \) satisfies the sequential-rationality requirement (1).

Case 1. Suppose that \( (\breve{a}^k_i, a^k_i, y^k_i) \notin D_{i,k} \). Then there exists \( l \leq k \) such that either \( \overline{A}_{-i} (\breve{a}_i(l), a_i(l), y_i(l)) \) is empty, or \( a_i(l) \neq \tilde{a}_i(l) \), and for all \( \tilde{y}_i(l) \in Y_i \), there exists \( \tilde{a}_{-i}(l) \in \overline{A}_{-i} (\breve{a}_i(l), a_i(l), y_i(l)) \) such that

\[
\text{supp } p_{-i} (\cdot | \tilde{a}_{-i}(l), a_i(l), y_i(l)) \setminus \text{supp } p_{-i} (\cdot | \breve{a}_{-i}(l), \tilde{a}_i(l), \tilde{y}_i(l)) \quad (43)
\]
is nonempty. That is, either player \( i \) has observed a deviation by the other players or has chosen a deviation that will be detected with positive probability by the other players. In this case, player \( i \) believes that his continuation payoff function, given the equilibrium reports, is

\[
S^i_\alpha \left( \tilde{a}^T, \tilde{\rho}^T \left( \tilde{a}^T, a^T, y^T \right) \right) = -d_1 + \Pi_i \left( \tilde{a}^T, \tilde{y}^T \right).
\]

Thus, it is optimal for player \( i \) to follow the recommended action in period \((k + 1)\).

Case 2. Suppose that \((\tilde{a}^i, a^i, y^i) \in D_{i,k}\). Since player \( i \) has not observed any deviations by the other players, he believes that \( \tilde{a}_{-i} = a_{-i} \). When the other players follow \( \alpha^T_{-i} \), by Part 1 of the lemma, for all \((\tilde{a}^T, a^T, y^T)\),

\[
\Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda^T, \alpha^T, \tilde{a}^k, a^k, y^k \right) = \Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda^T, \alpha^T, \tilde{a}^k, a^k, y^k \right).
\]

Hence his expected payoff conditional on \((\tilde{a}^i, a^i, y^i)\) becomes

\[
\sum_{\tilde{a}^T, a^T, y^T} \left( \sum_{l=1}^{T} \delta^{l-1} g_i (a(l)) + S^i_\alpha \left( \tilde{a}^T, \tilde{\rho}^T \left( \tilde{a}^T, a^T, y^T \right) \right) \right) \cdot \Pr \left( \tilde{a}^T, a^T, y^T \mid \lambda^T, \alpha^T, \tilde{a}^k, a^k, y^k \right) = \Pi_i \left( \alpha^T, \rho^T, S^T, \tilde{a}^k, a^k, y^k \right).
\]

The sequential-rationality requirement (1) is satisfied at \((\tilde{a}^i, a^i, y^i)\) because by definition \( \alpha^T_{-i} \) is a solution to

\[
\max_{a_{-i}(k)} \sum_{a_{-i}(k), y_i(k)} \Pi_i \left( \alpha^T_{-i}, \rho^T_{-i}, S^T, \tilde{a}^k, a^k, y^k \right) p_i (y_i(k) \mid \tilde{a}_{-i}(k), a_i(k)) \eta (\tilde{a}(k)).
\]

Next we show that the reporting strategy \( \rho^T_i \) satisfies the sequential-rationality requirement (1).

Case 1. Suppose that \((\tilde{a}^T, a^T, y^T) \in D_{i,T}\). Since player \( i \) has not observed any deviations by the other players, he believes that \( \tilde{a}^T_{-i} = a^T_{-i} \) and

\[
\tilde{\rho}^T_{-i} \left( \tilde{a}^T_{-i}, a^T_{-i}, y^T_{-i} \right) = \left( \tilde{a}^T_{-i}, y^T_{-i} \right).
\]

By the definition of \( S^i_\alpha \), if he reports any \( \tilde{a}_i \neq \tilde{a}_i \), then

\[
S^i_\alpha \left( \tilde{a}^T, \tilde{\rho}^T \left( \tilde{a}^T, a^T, y^T \right) \right) = -d_1 + \Pi_i \left( \tilde{a}^T, \tilde{y}^T \right).
\]
By the definition of $D_{i,T}$, there exists $\hat{y}_i^T \in Y_i^T$ such that if he reports $\hat{a}_i^T = \tilde{a}_i^T$ and $\hat{y}_i^T$, then his payoff is

$$\sum_{\tilde{a}_i^T, a_i^T, y_i^T} S_i^{\ast} (\tilde{a}_i^T, y_{-i}^T, \hat{y}_i^T) \Pr (\tilde{a}_i^T, a_i^T, y_i^T | \lambda^T \eta, \alpha_i^T, \alpha_{-i}^T, \tilde{a}_i^T, a_i^T, y_i^T).$$

By (37), it is optimal for player $i$ to report $\hat{a}_i^T = \tilde{a}_i^T$ and some $\hat{y}_i^T \in Y_i^T$. By definition $\hat{y}_i^T = \hat{p}_i^{T_1} (\hat{a}_i^T, a_i^T, y_i^T)$ is optimal.

Case 2. Suppose that there exists $k \leq T$ such that for all $a_{-i} \in A_{-i}$ with $\eta (a_{-i}, \tilde{a}_i(k)) > 0$, we have $p_i (y_i(k)|a_{-i}, a_i(k)) = 0$, i.e., $A_{-i} (\tilde{a}_i(k), a_i(k), y_i(k))$ is empty. Then player $i$ believes that some player $j$ has deviated from the recommended actions. Since player $j$ will report the deviation following $\hat{p}_j^{T_1}$, player $i$’s payoff function becomes

$$S_i^{\dagger} (\tilde{a}_i^T, \hat{p}_j^{T_1} (\tilde{a}_i^T, a_i^T, y_i^T)) = -d_1 + \Pi_i (\tilde{a}_i^T, \hat{y}_i^T).$$

Thus, it is optimal for player $i$ to report his actions and signals truthfully.

Case 3. Suppose that for all $k \leq T$ there exists $a_{-i} \in A_{-i}$ such that $\eta (a_{-i}, a_i(k)) > 0$ and $p_i (y_i(k)|a_{-i}, a_i(k)) > 0$, i.e., $A_{-i} (\tilde{a}_i(k), a_i(k), y_i(k))$ is nonempty; moreover, there exists $k \leq T$ such that $a_i(k) \neq \tilde{a}_i(k)$ and, for all $\hat{y}_i(k) \in Y_i$, there exists $\tilde{a}_{-i}(k) \in A_{-i} (\tilde{a}_i(k), a_i(k), y_i(k))$ and $y_{-i}(k)$ such that

$$y_{-i}(k) \in \text{supp} \ p_{-i} (\cdot | \tilde{a}_{-i}(k), a_i(k), y_i(k)) \setminus \text{supp} \ p_{-i} (\cdot | \tilde{a}_{-i}(k), \tilde{a}_i(k), \hat{y}_i(k)). \quad (44)$$

That is, player $i$ has not observed any deviations by the other players, he has deviated from the recommended action in some period, and for any misreport of private signals his deviation will be detected with positive probability. Let

$$\gamma_0 \equiv \min \left\{ p_{-i} (y_{-i}(k)|\tilde{a}_{-i}(k), a_i(k), y_i(k)) \eta (\tilde{a}_{-i}(k)|\tilde{a}_i(k)) | \right.$$

$$\left. y_{-i}(k) \in \text{supp} \ p_{-i} (\cdot | \tilde{a}_{-i}(k), a_i(k), y_i(k)) \setminus \text{supp} \ p_{-i} (\cdot | \tilde{a}_{-i}(k), \tilde{a}_i(k), \hat{y}_i(k)), \right.$$

$$\tilde{a}_{-i} \in A_{-i} (\tilde{a}_i(k), a_i(k), y_i(k)), (\tilde{a}_i(k), a_i(k), y_i(k)) \text{ satisfies (44)}, \hat{y}_i(k) \in Y_i, i \in N \right\}$$

denote the minimum of the conditional probabilities that a player’s deviation will be detected.
Since player $i$ has not observed any deviations by the other players, he believes that
\[ \tilde{\rho}^T_i (\tilde{a}^T_i, a^T_{-i}, y^T_{-i}) = (a^T_i, y^T_i). \]

If player $i$ reports $\hat{a}^T_i = a^T_i$, that is, he reports his deviation truthfully, then his payoff is
\[ -d_i + \Pi_i (\hat{a}^T, \hat{y}^T). \]

If player $i$ reports $\hat{a}^T_i = \tilde{a}^T_i$, that is, he does not report his deviation truthfully, for any $\hat{y}_i(k)$ his deviation will be discovered with probability at least $\gamma_0$. Hence his payoff is at most
\[ \Pi_i (1 - \gamma_0) - \delta^T v^*_i - g_i(a_N) \]

Choose $T_1$ such that for all $T \geq T_1$,
\[ (v^*_i - g_i(a_N)) \gamma_0 \geq \delta^*(T) - T(1 - \delta^*(T)) T \left( 2 \max_{i,a,y} w_i(a,y) - \min_{i,a,y} w_i(a,y) \right). \]

It follows that for all $T \geq T_1$ and $\hat{y}^T$,
\[ \Pi_i (1 - \gamma_0) - \delta^T v^*_i - g_i(a_N) \gamma_0 \geq -d_i + \Pi_i (\tilde{a}^T, \hat{y}^T), \]
and it is optimal for player $i$ to report truthfully. \qed

**B.2.3 The transfer functions and transition-probability function**

Using the transfer scheme $S^T$ for the $T$-period game, in the following we define a transfer function $\beta_{ij} : \Theta \times M \to \mathbb{R}_+$ to specify a payment $\beta_{ij}(\theta, m)$ from player $i$ to player $j$, for each pair of players $i, j \in \mathcal{N}$, and a transition-probability function $\zeta : \Theta \times M \to \mathbb{R}_+$ to specify a probability $\zeta(\theta, m)$ that the state in the next period is in the noncooperative state, for each profile of reports $(\theta, m) \in \Theta \times M$.

Proposition 3 is obviously true when $\sum_{i=1}^{n} g_i(\eta) + L^*(\eta) = \sum_{i=1}^{n} g_i(a_N)$. Henceforth, we assume that $\sum_{i=1}^{n} g_i(\eta) + L^*(\eta) > \sum_{i=1}^{n} g_i(a_N)$. Fix $\epsilon > 0$. By Lemma 3, we can pick $T_0$ so that when $T \geq T_0$ and $\delta \geq \delta^*(T)$,
\[ WL(T, \delta, S^{**}) < \min \left\{ |L^*(\eta)| + \epsilon, \sum_{i=1}^{n} g_i(\eta) - \sum_{i=1}^{n} g_i(a_N) \right\}. \]

(45)
Pick a vector ξ = (ξ_1, ξ_2, ..., ξ_n) with \( \sum_{i=1}^{n} \xi_i = 0 \) such that for each player \( i \)
\[
g_i(\eta) + \frac{1 - \delta}{1 - \delta} E_{\bar{a}^T, a^T, y^T} \left[ S_i^T (\bar{a}^T, a^T, y^T) \mid \lambda^{T\eta}, \alpha^{T\delta} \right] + \xi_i \geq g_i(a^N). \tag{46}
\]

If \((\theta, m) \notin A^T \times A^T \times Y^T\), or \((\theta, m) = (\bar{a}^T, \bar{a}^T, \bar{y}^T) \in A^T_i \times A^T_i \times Y^T_i, \bar{a}^T = \bar{a}^T\) and \(p(\bar{y}(\theta) | \bar{a}(\theta)) = 0\) for some \( k \leq T \), we set
\[
\beta_{ij}(\theta, m) = 0 \quad \text{for all } i, j \in \mathcal{N};
\]
\[
\zeta(\theta, m) = 1.
\]

From now on we assume that \((\theta, m) = (\bar{a}^T, \bar{a}^T, \bar{y}^T) \in A^T_i \times A^T_i \times Y^T_i, \bar{a}^T = \bar{a}^T\) and \(\Pr(\bar{y}(\theta) | \bar{a}^T) > 0\). Pick \(\delta \in (\delta^*(T), 1)\) such that for each player \( i \), each \((\bar{a}^T, a^T, y^T) \in A^T \times A^T \times Y^T\) and each \( \delta \geq \delta \),
\[
S_i^T (\bar{a}^T, a^T, y^T) + \frac{1 - \delta}{1 - \delta} \xi_i \geq \frac{\delta}{1 - \delta} \left( g_i(\eta) + \frac{1 - \delta}{1 - \delta} E_{\bar{a}^T, a^T, y^T} \left[ S_i^T (\bar{a}^T, a^T, y^T) \mid \lambda^{T\eta}, \alpha^{T\delta} \right] - g_i(a^N) \right). \tag{47}
\]

We define \(\zeta(\bar{a}^T, a^T, y^T)\) and \(\beta_{ij}(\bar{a}^T, a^T, y^T)\) as follows. First, for any \((\bar{a}^T, a^T, y^T) \in A^T \times A^T \times Y^T\) set
\[
\zeta(\bar{a}^T, a^T, y^T) = \frac{-\delta - (1 - \delta) \sum_{i=1}^{n} S_i^T (\bar{a}^T, a^T, y^T)}{\sum_{i=1}^{n} \left( g_i(\eta) + \frac{1 - \delta}{1 - \delta} E_{\bar{a}^T, a^T, y^T} \left[ S_i^T (\bar{a}^T, a^T, y^T) \mid \lambda^{T\eta}, \alpha^{T\delta} \right] - g_i(a^N) \right)} \tag{48}
\]

By (45) and (47), \(\zeta(\bar{a}^T, a^T, y^T) \in [0, 1]\) for all \((\bar{a}^T, a^T, y^T) \in A^T \times A^T \times Y^T\).

Next, set
\[
\beta_{ij}(\bar{a}^T, a^T, y^T) = \begin{cases} 
\beta_{ij}^{net}(\bar{a}^T, a^T, y^T) & \text{if } \beta_{ij}^{net}(\bar{a}^T, a^T, y^T) > 0, \\
0 & \text{otherwise}, \end{cases} \tag{49}
\]

where
\[
\beta_{ij}^{net}(\bar{a}^T, a^T, y^T) = \left(-S_i^T (\bar{a}^T, a^T, y^T) - \frac{1 - \delta}{1 - \delta} \xi_i \right) \delta^{-T+1}
\]
\[
-(1 - \delta)^{-1} \delta \zeta(\bar{a}^T, a^T, y^T) \left( g_i(\eta) + \frac{1 - \delta}{1 - \delta} E_{\bar{a}^T, a^T, y^T} \left[ S_i^T (\bar{a}^T, a^T, y^T) \mid \lambda^{T\eta}, \alpha^{T\delta} \right] + \xi_i - g_i(a^N) \right). \tag{50}
\]
When $\beta_i^{\text{net}}(\vec{a}^T, a^T, y^T) > 0$, $\sum_{j=1}^n \beta_{ij}(\vec{a}^T, a^T, y^T) = \beta_i^{\text{net}}(\vec{a}^T, a^T, y^T)$ and $\sum_{j=1}^n \beta_{ij}(\vec{a}^T, a^T, y^T) = 0$. When $\beta_i^{\text{net}}(\vec{a}^T, a^T, y^T) \leq 0$, $\sum_{j=1}^n \beta_{ij}(\vec{a}^T, a^T, y^T) = 0$ and $\sum_{j=1}^n \beta_{ij}(\vec{a}^T, a^T, y^T) = -\beta_i^{\text{net}}(\vec{a}^T, a^T, y^T)$; the last follows from the fact that, by (48) and (50), $\sum_{i=1}^n \beta_i^{\text{net}}(\vec{a}^T, y^T) = 0$ for all $(\vec{a}^T, a^T, y^T) \in A^T \times A^T \times Y^T$. Hence, for any player $i$ and any $(\vec{a}^T, a^T, y^T) \in A^T \times A^T \times Y^T$, the net side-payment received by player $i$ is always equal to

$$\sum_{j=1}^n (\beta_{ij}(\vec{a}^T, a^T, y^T) - \beta_{ij}(\vec{a}^T, a^T, y^T)) = -\beta_i^{\text{net}}(\vec{a}^T, a^T, y^T). \quad (51)$$

### B.2.4 Formal definition of the equilibrium strategy and belief

Given the perfect Bayesian equilibrium of the $T$-period game $\left(\alpha_i^{T\uparrow}, \rho_i^{T\uparrow}, \mu_i^{T\uparrow}\right)$ and the side-payment functions and transition-probability function $(\beta, \zeta)$, we can now write the trigger-strategy profile formally.

To simplify notation, we use $a_{i,l,k}$ to denote player $i$’s actions in the first $k$ periods of the $l$-th $T$-period block, i.e., $a_{i,l,k} = (a_{i,lT+1}, \ldots, a_{i,lT+k})$. Similarly, we let $\tilde{a}_{i,l,k} = (\tilde{a}_{i,lT+1}, \ldots, \tilde{a}_{i,lT+k})$ and $y_{i,l,k} = (y_{i,lT+1}, \ldots, y_{i,lT+k})$. For each $i \in \mathcal{N}$ and each $l \geq 0$, if $m_{i,(l+1)T} \in A_i^T \times Y_i^T$, then we write $m_{i,(l+1)T} = (\tilde{a}_{i,lT}^T, y_{i,lT}^T) \in A_i^T \times Y_i^T$.

Fix the following strategy for the mediator $(\lambda^n, \xi^n)$: in each period $t$ he always recommend actions according to the distribution $\eta$; if $t = (l+1)T$ for some $l \geq 0$, he reveals the recommended actions $\tilde{a}_i^{lT}$ in past $T$ periods; otherwise, he always reports nothing by sending the constant message $\emptyset$.

We say that a public history $h_{IT+k}^{\text{pub}}$, where $1 \leq k \leq T$, is in the cooperative state if (i) $l = 0$, or (ii) $h_{(l-1)T}^{\text{pub}}$ is in the cooperative state, $m_{i,(l-1)T} = (\tilde{a}_{i,(l-1)T}, \tilde{y}_{i,(l-1)T}) \in A_i^T \times Y_i^T$ for each $i$, either $\tilde{a}_{i,(l-1)T} \neq \tilde{a}_{i,(l-1)T}$ or $\tilde{a}_{i,(l-1)T} = \tilde{a}_{i,(l-1)T}$ and $\Pr(\tilde{y}_{i,(l-1)T}^{l-1,T} | \tilde{a}_{i,(l-1)T}^{l-1,T}) > 0$, $\tau_{ij,(l-1)T} = \beta_{ij}(\tilde{a}_{i,(l-1)T}^{l-1,T}, \tilde{a}_{j,(l-1)T}^{l-1,T}, \tilde{y}_{i,(l-1)T}^{l-1,T})$ for all $i, j$, and $\phi_{(l-1)T} \geq \zeta(\tilde{a}_{i,(l-1)T}^{l-1,T}, \tilde{a}_{j,(l-1)T}^{l-1,T}, \tilde{y}_{i,(l-1)T}^{l-1,T})$; otherwise, $h_{IT+k}^{\text{pub}}$ is in the non-cooperative state.

Let $C$ denote the set of public histories in the cooperative state.

Player $i$’s strategy $\sigma_i^* = (\alpha_i^*, \rho_i^*, b_i^*)$ is given as follows. For all $h_{IT+k}^{\text{pub}}$, $(\tilde{a}_{i,l,k}^T, \tilde{a}_{i,l,k}^T, y_{i,l,k}^T) \in C$
\( A_i^k \times A_i^k \times Y_i^k, \theta_{IT+k} \in \Theta, m_{IT+k} \in M, \) and all \( k \in \{1, 2, \ldots, T\} \), if \( h_{IT+k}^{\text{pub}} \in C \), then
\[
\begin{align*}
\alpha_i^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k-1}, y_i^{IT+k-1} \right) &= a_i^{T+1} \left( a_i^{l,k}, a_i^{l,k-1}, y_i^{l,k-1} \right), \\
\rho_i^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k}, y_i^{IT+k} \right) &= \begin{cases} \\
\rho_i^{T+1} \left( a_i^{l,k}, a_i^{l,k}, y_i^{l,k} \right) & \text{if } k = T \\
\emptyset & \text{if } 1 \leq k \leq T - 1 \\
\end{cases}, \\
b_{ij}^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k}, y_i^{IT+k}, \theta_{IT+k}, m_{IT+k} \right) = \begin{cases} \\
\beta_{ij} \left( \theta_{IT+k}, m_{IT+k} \right) & \text{if } k = T \text{ and } \theta_{IT+k} \in A^T \\
0 & \text{otherwise} \\
\end{cases}.
\end{align*}
\]

if \( h_{IT+k}^{\text{pub}} \notin C \), then
\[
\begin{align*}
\alpha_i^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k-1}, y_i^{IT+k-1} \right) &= a_i^N, \\
\rho_i^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k}, y_i^{IT+k} \right) &= \emptyset, \\
b_{ij}^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k}, y_i^{IT+k}, \theta_{IT+k}, m_{IT+k} \right) &= 0.
\end{align*}
\]

By Lemma 4, if \( h_{IT+k}^{\text{pub}} \in C \) and \( a_i^{IT,k-1} = a_i^{T,k-1} \), then
\[
\alpha_i^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k-1}, y_i^{IT+k-1} \right) = \tilde{a}_{i,IT+k}.
\] (52)

If \( h_{IT+k}^{\text{pub}} \notin C \), then
\[
\alpha_i^* \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k-1}, y_i^{IT+k-1} \right) = a_i^N. 
\] (53)

We define a system of beliefs \( \mu^* \) as follows.

For each \( \left( h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k}, y_i^{IT+k}, \theta_{IT+k}, m_{IT+k} \right) \in H_{i,IT+k}, i \in N, l \geq 0 \) and \( k \in \{1, 2, \ldots, T\} \), if \( h_{IT+k}^{\text{pub}} \in C \), let
\[
\begin{align*}
\mu_i^* \left( \tilde{a}_{-i}^{l,k-1}, a_i^{l,k-1}, y_i^{l,k-1} \right) h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k-1}, y_i^{IT+k-1} \right) &= \begin{cases} \\
\prod_{s=1}^{k-1} p_{-i} \left( y_{-i}(s), \tilde{a}_{-i}(s), \tilde{a}_i(s), y_i(s) \right) \prod_{s=1}^{k} \eta_{-i} \left( \tilde{a}_{-i}(s), \tilde{a}_i(s) \right) & \text{if } \tilde{a}_{-i}^{l,k-1} = a_{-i}^{l,k-1}; \\
0, & \text{otherwise}.
\end{cases} \\
\end{align*}
\]

and
\[
\begin{align*}
\mu_i^* \left( \tilde{a}_{-i}^{l,k}, a_i^{l,k}, y_i^{l,k} \right) h_{IT+k}^{\text{pub}}, a_i^{IT+k}, a_i^{IT+k}, y_i^{IT+k} \right) &= \begin{cases} \\
\prod_{s=1}^{k} p_{-i} \left( y_{-i}(s), \tilde{a}_{-i}(s), a_i(s), y_i(s) \right) \prod_{s=1}^{k} \eta_{-i} \left( \tilde{a}_{-i}(s), a_i(s) \right) & \text{if } \tilde{a}_{-i}^{l,k} = a_{-i}^{l,k}; \\
0, & \text{otherwise}.
\end{cases}
\end{align*}
\]
be beliefs of \( \tilde{a}_{l,k}^{l_i}, a_{l,k}^{l_i}, y_{l,k}^{l_i} \) derived by Bayes’ rule and \( \alpha^* \). The beliefs about the information before the current block, \( \tilde{a}_{l,i}, a_{l,i}, y_{l,i} \), the beliefs in the noncooperative state will be irrelevant in checking the sequential-rationality condition.

**B.2.5 Theorem 1**

We claim that, given \( (\lambda^\eta, \xi^*) \), the trigger-strategy profile \( \sigma^* \) and the belief system \( \mu^* \) constitute a perfect Bayesian equilibrium. It is obvious that the continuation strategies in the noncooperative state are self-enforcing, and the belief system \( \mu^* \) satisfies Bayes’ rule. It remains to show that \( \sigma^* \) satisfies the sequential-rationality requirement (1) in the cooperative state.

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. Then, given the report profile \( (\tilde{a}^{l,T}, \tilde{a}^{l,T}, y^{l,T}) \), by (51) the discounted continuation payoff of player \( i \) at the end of the \( T \)-period block before side-payments are made is

\[
-\beta_i^{\text{net}} (\tilde{a}^{l,T}, \tilde{a}^{l,T}, y^{l,T}) + (\zeta (\tilde{a}^{l,T}, \tilde{a}^{l,T}, y^{l,T}) (v_i^N - v_i^*) + v_i^*) \frac{\delta}{1 - \delta}.
\]

On the equilibrium paths,

\[
\alpha^* \left( h^{\text{pub}}_{l_iT+k}, a_{i}^{lT+k-1}, y_{i}^{lT+k-1} \right) = \alpha^T \left( \tilde{a}_i^{lT}, a_{i}^{lT-1}, y_{i}^{lT-1} \right)
\]

and

\[
\rho^* \left( h^{\text{pub}}_{l_iT+T}, \tilde{a}_{i}^{lT+T}, a_{i}^{lT+T}, y_{i}^{lT+T} \right) = \left( a_i^{lT}, y_i^{lT} \right).
\]

Hence, by standard arguments,

\[
v_i^* = (1 - \delta^T) g_i(\eta) - E_{a^{l,T},a^{l,T},y^{l,T}} [\beta_i^{\text{net}} (\tilde{a}^{l,T}, a^{l,T}, y^{l,T}) | \lambda^\eta, \alpha^{T*}] \delta^{T-1} (1 - \delta) + (E_{a^{l,T},a^{l,T},y^{l,T}} [\zeta (\tilde{a}^{l,T}, a^{l,T}, y^{l,T}) | \lambda^\eta, \alpha^{T*}] (v_i^N - v_i^*) + v_i^*) \delta^T.
\]

Substituting (50) into (55) and rearranging terms, we have

\[
v_i^* = g_i(\eta) + \frac{1 - \delta}{1 - \delta^T} E_{a^{l,T},a^{l,T},y^{l,T}} [S_i^T (\tilde{a}^{l,T}, a^{l,T}, y^{l,T}) | \lambda^\eta, \alpha^{T*}] + \xi_i
\]

\[
= g_i(\eta) + \frac{1 - \delta}{1 - \delta^T} E_{a^{l,T},a^{l,T},y^{l,T}} [S_i^{T*} (\tilde{a}^{l,T}, y^{l,T}) | \lambda^\eta, \alpha^{T*}] + \xi_i.
\]
It is straightforward to see that

\[ \nu_i^N = g_i(a^N). \]

Now, by (50) and (56), (54) is equivalent to

\[
\left( S_i^T \left( \tilde{\alpha}_T, \tilde{\alpha}_T^*, y_T \right) + \frac{1 - \delta^T}{1 - \delta} \xi_i \right) \delta^{-T+1} + \nu_i^* \frac{\delta}{1 - \delta}.
\]

(57)

According to the trigger-strategy profile, the players will switch to the non-cooperative state with probability one if any player fails to make the required side-payments. It follows from (47) that when \( \delta \) 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 \( \alpha_i^T \) and a reporting strategy \( \rho_i^T \), at the beginning of the cooperative state, is equal to

\[
U_i^T \left( \alpha_i^T, \rho_i^T, \alpha_{-i}^T, \rho_{-i}^T; S^T \right) + \frac{1 - \delta^T}{1 - \delta} \xi_i + \nu_i^* \frac{\delta^T}{1 - \delta}.
\]

It follows from Lemma 4 that, when \( T \) is sufficiently large and \( \delta \) 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 \( \delta \) 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

\[
g_i(\eta) + \frac{1 - \delta}{1 - \delta^T} E_{\tilde{\alpha}, \alpha, y} \left[ S_i^* \left( \tilde{\alpha}_T^*, y_T \right) \mid \lambda^T, \alpha^T \right] + \xi_i.
\]

It follows from (45) that the total expected equilibrium payoff is greater than

\[ \sum_{i=1}^{n} g_i(\eta) + L^*(\eta) - \epsilon. \]
C Proof of Proposition 1

Let

\[ \Pr(y|a_i, \tilde{a}_{-i}, \rho_i) = \sum_{y_{-i}' \in \rho_i^{-1}(y)} p(y_{-i}', y_i'|a_i, \tilde{a}_{-i}) \]

\[ \Pr(\tilde{a}_i', \tilde{a}_{-i}, y|\eta, \tilde{a}_i, \rho_i) = \Pr(y|a_i, \tilde{a}_{-i}, \rho_i) \eta(\tilde{a}_i', \tilde{a}_{-i}) I_{\tilde{a}_i}(\tilde{a}_i') \]

\[ \Pr(\tilde{a}_i', \tilde{a}_{-i}, y)\eta, \tilde{a}_i) = \Pr(y|\tilde{a}) \eta(\tilde{a}_i', \tilde{a}_{-i}) I_{\tilde{a}_i}(\tilde{a}_i') \]

where

\[
I_{\tilde{a}_i}(\tilde{a}_i') = \begin{cases} 1, & \text{if } \tilde{a}_i = \tilde{a}_i' \\ 0, & \text{otherwise.} \end{cases}
\]

Then

\[
\pi^*_{a_i}(\tilde{a}_{-i}, \tilde{a}_i', y_{-i}, y_i) = \eta(\tilde{a}_{-i}, a_i') \sum_{a_i, \rho_i} s_i(\tilde{a}_i', a_i, \rho_i) \sum_{y_{-i}' \in \rho_i^{-1}(y)} p(y_{-i}', y_i'|a_i, \tilde{a}_{-i})
\]

\[ = \sum_{a_i, \rho_i} s_i(\tilde{a}_i, a_i, \rho_i) \Pr(y|a_i, \tilde{a}_{-i}, \rho_i) \eta(\tilde{a}_{-i}, \tilde{a}_i') \]

\[ = \sum_{\tilde{a}_i \in A, a_i, \rho_i} \sum_{a_i, \rho_i} s_i(\tilde{a}_i, a_i, \rho_i) \Pr(y|a_i, \tilde{a}_{-i}, \rho_i) \eta(\tilde{a}_{-i}, \tilde{a}_i') I_{\tilde{a}_i}(\tilde{a}_i')
\]

Recall:

\[
L^*(\eta) = \max_{x: A \times Y \times Y \to \mathbb{R}} \sum_{i=1}^{n} \sum_{(a,y) \in A \times Y} x_i(a, y) \Pr(a, y|\eta)
\]

subject to

\[
\sum_{\tilde{a}_{-i} \in A_{-i}} \sum_{y \in Y} \left[ \Pr(y|a_i, \tilde{a}_{-i}, \rho_i) \eta(\tilde{a}) - p(y|\tilde{a}) \eta(\tilde{a}) \right] x_i(\tilde{a}, y)
\]

\[ \leq \sum_{\tilde{a}_{-i} \in A_{-i}} (g_i(\tilde{a}) - g_i(\tilde{a}_{-i}, a_i)) \eta(\tilde{a}) \quad \text{for all } \tilde{a}_i, a_i, \rho_i, i \]

and

\[
\sum_{i=1}^{n} \sum_{(a,y) \in A \times Y} x_i(a, y) \Pr(a, y|\eta, \omega) \leq 0 \quad \text{for all } \omega \in P.
\]
Rewrite (59) as
\[ \sum_{(\tilde{a}_i',\tilde{a}_{-i},y) \in A \times Y} \left[ \Pr ( y | a_i, \tilde{a}_{-i}, \rho_i) \eta (\tilde{a}_i', \tilde{a}_{-i}) I_{\tilde{a}_i} (\tilde{a}_i') - p ( y | \tilde{a}) \eta (\tilde{a}_i', \tilde{a}_{-i}) I_{\tilde{a}_i} (\tilde{a}_i') \right] x_i (\tilde{a}_i', \tilde{a}_{-i}, y) \]
\begin{align*}
&\leq \sum_{\tilde{a}_{-i} \in A_{-i}} (g_i (\tilde{a}) - g_i (\tilde{a}_{-i}, a_i)) \eta (\tilde{a}) \quad \text{for all } \tilde{a}_i, a_i, \rho_i, i, \quad (61)
\end{align*}
or,
\[ \sum_{(\tilde{a}_i',\tilde{a}_{-i},y) \in A \times Y} \left[ \Pr ( y | a_i, \tilde{a}_{-i}, y | \eta, \tilde{a}_i, a_i, \rho_i) - \Pr ( y | \eta, \tilde{a}_i, a_i) \right] x_i (\tilde{a}_i', \tilde{a}_{-i}, y | \tilde{a}_i) \\
&\leq \sum_{\tilde{a}_{-i} \in A_{-i}} (g_i (\tilde{a}) - g_i (\tilde{a}_{-i}, a_i)) \eta (\tilde{a}) \quad \text{for all } \tilde{a}_i, a_i, \rho_i, i, \quad (62)
\]

The dual program is to choose \( \lambda_i (\tilde{a}_i, a_i, \rho_i) \geq 0 \) for each \( (\tilde{a}_i, a_i, \rho_i) \) and \( i \), and \( \nu(\omega) \geq 0 \) for each \( \omega \in \mathcal{P} \) to maximize the total weighted gains from deviation
\[
\sum_{i=1}^{n} \sum_{\tilde{a}_i, a_i, \rho_i} \lambda_i (\tilde{a}_i, a_i, \rho_i) \sum_{\tilde{a}_{-i} \in A_{-i}} (g_i (\tilde{a}) - g_i (\tilde{a}_{-i}, a_i)) \eta (\tilde{a}),
\]
subject to
\[
\sum_{\tilde{a}_i, a_i, \rho_i} \lambda_i (\tilde{a}_i, a_i, \rho_i) (\Pr (\cdot | \eta, \tilde{a}_i, a_i, \rho_i) - \Pr (\cdot | \eta, \tilde{a}_i)) + \sum_{\omega \in \mathcal{P}} \nu(\omega) \Pr (\cdot | \eta, \omega) = \Pr (\cdot | \eta) \quad \text{for each } i. \quad (63)
\]
Here, each multiplier \( \lambda_i (\tilde{a}_i, a_i, \rho_i) \) corresponds to the incentive constraint (59) for player \( i \) and the deviation \( (\tilde{a}_i, a_i, \rho_i) \), and each \( \nu(\omega) \) corresponds to the budget-balance constraint (60) in state \( \omega \in \mathcal{P} \). Since the dual program is feasible with \( \lambda_i (\tilde{a}_i, a_i, \rho_i) = 0 \) for each \( i \) and each \( (\tilde{a}_i, a_i, \rho_i) \), and \( \nu(\omega) = \Pr (\omega | \eta) \) for each \( \omega \in \mathcal{P} \), by the strong duality theorem, the value of the dual program is \( L^*(\eta) \), i.e.,
\[
L^*(\eta) = \max_{\lambda \geq 0, \nu \geq 0} \sum_{i=1}^{n} \sum_{\tilde{a}_i, a_i, \rho_i} \lambda_i (\tilde{a}_i, a_i, \rho_i) \sum_{\tilde{a}_{-i} \in A_{-i}} (g_i (\tilde{a}) - g_i (\tilde{a}_{-i}, a_i)) \eta (\tilde{a}) \quad (64)
\]
subject to (63).

Let
\[
L = \max_{(s_1, \ldots, s_n)} \gamma(s_1) \sum_{i=1}^{n} \sum_{\tilde{a}_i, a_i, \rho_i} s_i (\tilde{a}_i, a_i, \rho_i) \sum_{\tilde{a}_{-i} \in A_{-i}} (g_i (\tilde{a}) - g_i (\tilde{a}_{-i}, a_i)) \eta (\tilde{a}) \quad (65)
\]
subject to: \( \pi^{s_1} = \cdots = \pi^{s_n} \in Q(\eta) \).
We want to show that $L^*(\eta) = \mathcal{L}$.

We first show that $L^*(\eta) \leq \mathcal{L}$. Since $\mathcal{L} \geq 0$, this holds if $L^*(\eta) = 0$. Suppose that $L^*(\eta) > 0$ and $(\lambda, \nu)$ is a solution to Program (64). We claim that $\nu(\omega) = 0$ for some $\omega$. Otherwise, by complementary slackness, the constraint (60) holds with equality for all $\omega$. Hence $L^*(\eta) = 0$, a contradiction.

For each player $i$, fix some $\tilde{a}_i^* \in A_i$. Set

$$\lambda_0 = \max_{\tilde{a}_i, a_i, \rho_i} \lambda_i (\tilde{a}_i, a_i, \rho_i);$$

$$s_i (\tilde{a}_i, a_i, \rho_i) = \frac{1}{\lambda_0} \lambda_i (\tilde{a}_i, a_i, \rho_i) \text{ for each } (\tilde{a}_i, a_i, \rho_i) \neq (\tilde{a}_i^*, a_i^*, \rho_i^*);$$

$$s_i (\tilde{a}_i^*, a_i^*, \rho_i^*) = 1 - \frac{1}{\lambda_0} \sum_{(\tilde{a}_i, a_i, \rho_i) \neq (\tilde{a}_i^*, a_i^*, \rho_i^*)} \lambda_i (\tilde{a}_i, a_i, \rho_i);$$

$$\tilde{\nu} (\omega) = \frac{1}{\lambda_0} \left[ \lambda_0 + 1 \right] \Pr (\omega | \eta) - \nu (\omega) \right] \text{ for each } \omega.$$

Then we can write (63) as

$$\pi^{s_1} = \cdots = \pi^{s_n} = \sum_{\omega \in P} \tilde{\nu} (\omega) \Pr (\cdot | \eta, \omega) \in Q(\eta).$$

Moreover, since $\nu (\omega) = 0$ for some $\omega$,

$$\gamma(s_1) = \left( \max_{\omega \in P} \frac{\tilde{\nu} (\omega)}{\Pr (\omega | \eta)} - 1 \right)^{-1} = \left( \max_{\omega \in P} \frac{\lambda_0 + 1}{\lambda_0} \Pr (\omega | \eta) - \nu (\omega) \right) \Pr (\omega | \eta) - 1 \right)^{-1}$$

$$= \lambda_0 \left( 1 - \min_{\omega \in P} \frac{\nu (\omega)}{\Pr (\omega | \eta)} \right)^{-1} = \lambda_0.$$

Hence $s$ is a candidate for Program (65) with value $L^*(\eta)$. It follows that $L^*(\eta) \leq \mathcal{L}$.

Conversely, suppose that $s$ is a solution to Program (65). Hence

$$\pi^{s_1} = \cdots = \pi^{s_n} = \sum_{\omega \in P} \tilde{\nu} (\omega) \Pr (\cdot | \eta, \omega) \in Q(\eta)$$

(66)

for some $\tilde{\nu}$. By rearranging terms we can write (66) as

$$\gamma(s_1) \sum_{\tilde{a}_i, a_i, \rho_i} s_i (\tilde{a}_i, a_i, \rho_i) \left( \Pr (\cdot | \eta, \tilde{a}_i, a_i, \rho_i) - \Pr (\cdot | \eta, \tilde{a}_i) \right)$$

$$+ \sum_{\omega} \left[ (1 + \gamma(s_1)) \Pr (\omega | \eta) - \gamma(s_1) \tilde{\nu} (\omega) \right] \Pr (\cdot | \eta) = \Pr (\cdot | \eta) \text{ for each } i,$$
By the definition of $\gamma(s_1)$ we have

$$(1 + \gamma(s_1)) \Pr(\omega|\eta) - \gamma(s_1)\nu(\omega) \geq 0 \text{ for each } \omega.$$ 

Hence

$$\lambda_i(\tilde{a}_i, a_i, \rho_i) = \gamma(s_1)s_i(\tilde{a}_i, a_i, \rho_i) \text{ for each } (\tilde{a}_i, a_i, \rho_i) \text{ and each } i \in \mathcal{N};$$

$$\nu(\omega) = (1 + \gamma(s_1)) \Pr(\omega|\eta) - \gamma(s_1)\tilde{\nu}(\omega) \text{ for each } \omega$$

is a candidate for Program (64) with value $\mathcal{L}$. It follows that $\mathcal{L} \leq L^*(\eta)$.

\section{Proof of Theorem 2}

Suppose $\eta$ is strictly enforceable. For any $\epsilon > 0$, we define a correlated distribution $\eta_\epsilon$ that converges to $\eta$ as $\epsilon$ tends to 0. For any player $i$, any $a_{-i}$ in the support of $\eta_{-i}$, and any $y_{-i} \in Y_{-i}$, let $f_i(a_{-i}, y_{-i})$ be an action in $A_i$ such that

1. $f_i(a_{-i}, y_{-i}) \in \arg\max_{a_i'} p_{-i}(y_{-i}|a_i', a_{-i})$,

2. $p(\cdot|f_i(a_{-i}, y_{-i}), a_{-i})$ is an extremal point in the set

$$\left\{ p(\cdot|a_i, a_{-i}) \middle| a_i \in \arg\max_{a_i'} p_{-i}(y_{-i}|a_i', a_{-i}) \right\},$$

and (3) $f_i(a_{-i}, y_{-i})$ maximizes $g_i(a_i', a_{-i})$ over $a_i'$ that satisfies (1) and (2). Let $k$ be the cardinality of the set

$$\mathcal{A} \equiv \{(f_i(a_{-i}, y_{-i}), a_{-i})| (a_{-i}, y_{-i}) \in \text{support } (\eta_{-i}) \times Y_{-i}, i \in \mathcal{N}\}.$$ 

For any action profile $a \in A$, we let $\gamma_a \in \Delta(A)$ denote the distribution that assigns probability one to $a$. That is, $\gamma_a(a) = 1$ and $\gamma_a(a') = 0$ for all $a' \neq a$. Define

$$\eta_\epsilon \equiv (1 - k\epsilon) \eta + \epsilon \sum_{a \in \mathcal{A}} \gamma_a.$$ 

We first show that $\eta_\epsilon$ is strictly enforceable, when $\epsilon$ is sufficiently small. For any player $i$, if he receives a recommendation $a_i$ under $\eta_\epsilon$, then either $a_i$ is in the
support of $\eta_i$, or it is equal to $f_i(a_{-i}, y_{-i})$ for some $a_{-i}$ in the support of $\eta_{-i}$ and some $y_{-i} \in Y_{-i}$. If $a_i$ is in the support of $\eta_i$, since $\eta$ is strictly enforceable, any deviation will either result in a different distribution against $\eta$ or a strictly lower payoff for $i$. If a deviation from $a_i$ results in a different distribution against $\eta$, it will lead to a different distribution against $\eta$. If a deviation leads to a strictly lower payoff against $\eta$, it will also lead to a strictly lower payoff against $\eta$ when $\epsilon$ is sufficiently small. Lastly, by construction any $f_i(a_{-i}, y_{-i})$ is strictly enforceable.

Let

$$\mu_\epsilon(a, y) \equiv p(y|a) \eta_\epsilon(a), \quad \text{for all } (a, y) \in A \times Y$$

denote the distribution generated by $\eta_\epsilon$ and $p$, and let $P_\epsilon$ denote the meet of players’ information partitions induced by $\mu_\epsilon$. Note that $P_\epsilon$ could be different from the meet induced by $\eta$ and $p$. Let

$$Q(\eta_\epsilon) \equiv co\{\mu_\epsilon(\cdot|\omega) | \omega \in P_\epsilon\}.$$

Next we show that, for any player $i$ and any strategy $s_i$,

$$\pi^{s_i} \in Q(\eta_\epsilon) \quad \text{implies that} \quad \pi^{s_i} = \mu_\epsilon.$$

Suppose, towards a contradiction, that $\pi^{s_i} \in Q(\eta_\epsilon)$ but $\pi^{s_i} \neq \mu_\epsilon$. Then there exists $(a, y)$ in the support of $\mu_\epsilon$ such that

$$\pi^{s_i}(a, y) > \mu_\epsilon(a, y).$$
By the definition of $f_i (a_{-i}, y_{-i})$,

$$\sum_{y'_i} \pi^{s_i} (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i})$$

$$= \sum_{y'_i} \eta_e (f_i (a_{-i}, y_{-i}), a_{-i}) \sum_{a'_{i}, \rho_i} s_i (f_i (a_{-i}, y_{-i}), a'_{i}, \rho_i) \sum_{y''_i \in \rho_i^{-1} (y'_i)} p (y''_i, y_{-i} | a'_{i}, a_{-i})$$

$$= \eta_e (f_i (a_{-i}, y_{-i}), a_{-i}) \sum_{a'_{i}, \rho_i} s_i (f_i (a_{-i}, y_{-i}), a'_{i}, \rho_i) \sum_{y''_i \in \rho_i^{-1} (y'_i)} p (y''_i, y_{-i} | a'_{i}, a_{-i})$$

$$= \eta_e (f_i (a_{-i}, y_{-i}), a_{-i}) \sum_{a'_{i}, \rho_i} s_i (f_i (a_{-i}, y_{-i}), a'_{i}, \rho_i) \sum_{y''_i \in \rho_i^{-1} (y'_i)} p (y''_i, y_{-i} | a'_{i}, a_{-i})$$

$$\leq \eta_e \left( f_i (a_{-i}, y_{-i}), a_{-i} \right) p_{-i} (y_{-i} | f_i (a_{-i}, y_{-i}), a_{-i})$$

$$= \eta_e \left( f_i (a_{-i}, y_{-i}), a_{-i} \right) \sum_{y'_i} \mu_e (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i})$$

Hence

$$\pi^{s_i} (a, y) \sum_{y'_i} \pi^{s_i} (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i}) > \sum_{y'_i} \mu_e (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i}).$$

Since $\pi^{s_i} (a, y) > 0$ and $\pi^{s_i} \in Q (\eta_e)$, there exists $\omega \in P_\epsilon$ such that $\mu_e (a, y | \omega) > 0$.

Similarly, for each $y'_i$ such that $\pi^{s_i} (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i}) > 0$, since $(f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i})$ and $(a, y)$ have the same $-i$ component, $\mu_e (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i} | \omega) > 0$ for the same $\omega$. It follows that

$$\frac{\pi^{s_i} (a, y | \omega)}{\sum_{y'_i} \pi^{s_i} (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i} | \omega)} = \frac{\pi^{s_i} (a, y)}{\sum_{y'_i} \pi^{s_i} (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i})}$$

$$> \frac{\mu_e (a, y)}{\sum_{y'_i} \mu_e (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i})}$$

$$= \frac{\mu_e (a, y | \omega)}{\sum_{y'_i} \mu_e (f_i (a_{-i}, y_{-i}), a_{-i}, y'_i, y_{-i} | \omega)},$$

which contradicts the supposition that $\pi^{s_i}$ belongs to $Q (\eta_e)$.

We have established that if $\pi^{s_i} \in Q (\eta_e)$, then $\pi^{s_i} = \mu_e$. But $\eta_e$ is strictly enforceable, when $\epsilon$ is small. Hence, if $s_i$ is a deviating strategy and $\pi^{s_i} = \mu_e$, player $i$ must be strictly worse off for choosing $s_i$.  

53
References


