The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
23-018

Uniformly Strict Equilibrium for Repeated Games with Private Monitoring and Communication

RICHARD MCLEAN
Rutgers University

ICHIRO OBARA
UCLA

ANDREW POSTLEWAITE
University of Pennsylvania

October 2023
revised February 5, 2024

Uniformly Strict Equilibrium for Repeated Games with Private Monitoring and Communication*

Richard McLean
Rutgers University

Ichiro Obara
UCLA

Andrew Postlewaite
University of Pennsylvania

February 5, 2024

Abstract

Cooperation through repetition is an important theme in game theory. In this regard, various celebrated "folk theorems" have been proposed for repeated games in increasingly more complex environments. There has, however, been insufficient attention paid to the robustness of a large set of equilibria that is needed for such folk theorems. Starting with perfect public equilibrium as our starting point, we study uniformly strict equilibria in repeated games with private monitoring and direct communication (cheap talk). We characterize the limit equilibrium payoff set and identify the conditions for the folk theorem to hold with uniformly strict equilibrium.

Keywords: Cheap talk, Communication, Folk theorem, Private monitoring, Repeated games, Robustness, Strict equilibrium

JEL Classifications: C72, C73, D82

*We thank George Mailath for very helpful discussions and the audience in numerous conferences and seminars for helpful comments. Postlewaite gratefully acknowledges support from National Science Foundation grant SBR-9810693.

1 Introduction

Cooperation through repetition is an important theme in game theory. In this regard, various celebrated "folk theorems" have been proposed for repeated games in increasingly more complex environments. There has, however, been insufficient attention paid to the robustness of a large set of equilibria that is needed for such folk theorems.

In this paper, we study *uniformly strict equilibria* in repeated games with private monitoring and direct communication (cheap talk). Our starting point is perfect public equilibrium (PPE) ([9]). In each period, players take actions simultaneously, observe private signals, and send public messages simultaneously. A perfect public equilibrium is a profile of public strategies that specifies a Nash equilibrium as their continuation play after every public history (a sequence of past message profiles). We impose strict incentives at every public history by requiring that, in each period, a player would incur a positive payoff loss (in terms of the value at the period) when deviating in either action or message from the equilibrium strategy. We also require such payoff losses from a unilateral deviation to be uniformly bounded away from 0 across all public histories.

It is well known that strict equilibrium has desirable robustness properties. For example, strict equilibria survive most equilibrium refinements in strategic form games. In our setting of infinitely repeated games, our uniform strictness requirement is a natural strengthening of strict equilibrium.

We present two main results. Our first result is a characterization of the limit set of uniformly strict perfect public equilibrium payoffs via a collection of static programming problems. We follow the approach of Fudenberg and Levine [7] (henceforth FL) to characterize the limit equilibrium payoff set. It also builds on other classic results from Abreu, Pearce and Stacchetti [1] and Fudenberg, Levine and Maskin [9]. We adapt their ideas to our model and generalize them by introducing uniformly strict incentives. In our second result, we establish a folk theorem by identifying conditions ensuring that this limit set coincides with the set of feasible and individually rational payoffs generated by the data of the underlying stage game.

There is a large literature dealing with folk theorems for repeated games with varying assumptions regarding public or private monitoring with or without communication. Most relevant to our paper are the various folk theorems for repeated games with private monitoring and communication ([2], [3], [5], [8], [12], [13], [14], [18], [23]).[1]

[1]There is an extensive literature on folk theorems for repeated games with private mon-

Our detectability and identifiability conditions for the folk theorem are similar to and weaker than the conditions (A2) and (A3) in Kandori and Matsushima [12]. (A2) and (A3) imply that, for any pair of players, their deviations are detectable and identifiable (i.e. one player's deviation can be statistically distinguished by the other player's deviation) based on the private signals of the other $n-2$ players. Our detectability and identifiability condition instead impose a similar restriction on the joint distributions of the messages of all players. Their conditions allow each player's future payoff independent of her message. This indifference makes truth-telling incentive compatible for each player. On the other hand, we require uniform strictness of incentive for sending any (nontrivial) message.[2]

Our conditions are also similar to the sufficient conditions in Tomala [23], but Tomala studies a type of perfect equilibrium with mediated communication, which is more flexible than cheap talk, and does not impose strict incentive constraints. As a consequence, the conditions for the folk theorem in Tomala [23] are weaker than ours.

2 Preliminaries

2.1 Repeated Games with Private Monitoring and Communication

Stage Game

We present the model of repeated games with private monitoring and communication. The set of players is $N = \{1, ..., n\}$. The game proceeds in stages and in each stage t, player i chooses an action from a finite set A_i. An action profile is denoted by $a = (a_1, ..., a_n) \in \Pi_i A_i := A$. Stage game payoffs are given by $g : A \to \mathbb{R}^n$. We denote the resulting stage game by $G = (N, A, g)$. Actions are not publicly observable. Instead, each player i observes a private signal s_i from a finite set S_i. A private signal profile is denoted $s = (s_1, .., s_n) \in \Pi_i S_i := S$. For each $a \in A$, $p(\cdot|a) \in \Delta(S)$ is the distribution on S given action profile a. We assume that the marginal distributions have

itoring and without communication, including [4], [6] , [11], [15], [16], [17], [19], [21], [22]. They usually rely on non-strict equilibrium (such as belief-free equilibrium) to establish the folk theorem.

[2] [12] also discusses a way to provide strict incentive for truth-telling via a scoring rule, but the strict incentive vanishes in the limit for the minmax points for their folk theorem (Theorem 2). We instead fix the level of the required strict incentive first, then prove the folk theorem by letting $\delta \to 0$.

full support, that is, $\sum_{s_{-i}} p\left(s_i, s_{-i}|a\right) > 0$ for all $s_i \in S_i$, $a \in A$ and $i \in N$.

Players communicate publicly each period. Player i sends a public message m_i from a finite set M_i after observing a private signal s_i in each period.[3] Player i's *message strategy* $\rho_i : S_i \to M_i$ in the stage game is a mapping from private signals to public messages. Let R_i be the set of player i's message strategies. An action profile $a \in A$ and a profile of message strategies $\rho = (\rho_1, ..., \rho_n) \in \Pi_i R_i := R$ generates a distribution $\tilde{p}(\cdot|(a, \rho))$ over public messages $M = \Pi_i M_i$, where

$$\tilde{p}(m|(a, \rho)) := \sum_{s \in S: \rho_i(s_i) = m_i, \forall i} p(s|a).$$

We normalize payoffs so that each player's *pure strategy* minmax payoff is 0 in the stage game. The pure strategy minmax payoff is the relevant payoff lower bound for our folk theorem because we study equilibrium with strict incentives and without any mediator. Note that the pure strategy minmax payoff may be strictly larger than the mixed minmax payoff. The set of feasible payoff profiles is $V(G) = co\left\{g\left(a\right)|a \in A\right\}$. Let $A(G) \subseteq A$ be the set of action profiles that generate an extreme point in $V(G)$. Finally $V^*(G) = \{v \in V(G)| v \geq 0\}$ is the set of feasible, individually rational payoff profiles.

Repeated Game with Public Communication

In the repeated game, play proceeds in the following way. At the beginning of period $t \geq 1$, player i chooses an action contingent on (h_i^t, h^t), where $h_i^t \in H_i^t = A_i^{t-1} \times S_i^{t-1}$ is player i's private history that consists of her private actions and private signals and $h^t \in H^t = M^{t-1}$ is the public history of message profiles.[4] Player i also chooses a message strategy $\rho_i \in R_i$ contingent on (h_i^t, h^t, a_i). Then player i's pure strategy $\sigma_i = (\sigma_i^a, \sigma_i^m)$ consists of an "action" component $\sigma_i^a : \bigcup_t [H_i^t \times H^t] \longrightarrow A_i$ and a "message" component $\sigma_i^m : \bigcup_t [H_i^t \times H^t \times A_i] \longrightarrow R_i$.

A strategy σ_i is a public strategy if in any period t both $\sigma_i^a(h_i^t, h^t)$ and $\sigma_i^m(h_i^t, h^t, a_i)$ are independent of private history $h_i^t \in H_i^t$. For the sake of simple exposition, we drop h_i^t from any public strategies. We denote player i's action and on-path message strategy at h^t for public strategy σ_i by $\sigma_i(h^t) = (\sigma_i^a(h^t), \sigma_i^m(h^t, \sigma_i^a(h^t))) \in A_i \times R_i$. A pure strategy profile σ induces a probability measure on A^∞. Player i's discounted average payoff given a profile of strategies $\sigma = (\sigma_1, ..., \sigma_n)$ is $(1-\delta)E[\sum_{t=1}^{\infty} \delta^{t-1} g_i(\tilde{a}^t)|\sigma]$, where the expectation is taken with respect to this measure.

[3] We can support the largest set of equilibria by using $M_i = S_i$. But we use a more general message space M_i to allow for the possibility of restricted message spaces.

[4] We define $H^1 = H_i^1 = \{\varnothing\}$ for all $i \in N$.

2.2 Uniformly Strict Perfect Public Equilibrium

A profile of public strategies σ is a perfect public equilibrium if its continuation strategies constitute a Nash equilibrium after every public history([9]). In this paper, we impose an additional robustness requirement by requiring uniformly strict incentive compatibility at every public history. Let $w^{\sigma}(h^t)$ be a profile of discounted average continuation payoffs at public history $h^t \in H$ given a public strategy profile σ. Given σ and h^t, let Σ_i^{σ,h^t} be the set of deviations $(a_i', \rho_i') \in A_i \times R_i$ such that $a_i' \neq \sigma_i^a(h^t)$ or i's unilateral deviation from $\sigma_i(h^t)$ to (a_i', ρ_i') changes the distribution of continuation payoff profiles $w^{\sigma}(h^t, \cdot)$ from period $t+1$. We call such one-shot deviations *nontrivial deviations* at h^t with respect to w^{σ}. Any other one-shot deviation is called a *trivial deviation*, as it does not change any outcome in the current period and in the future at all.

We define *η-uniformly strict perfect public equilibrium* (η-USPPE) as follows.

Definition 1 *(η-USPPE) A profile of public strategies σ is an η -uniformly strict perfect public equilibrium for $\eta \geq 0$ if the following conditions are satisfied for any $h^t \in H$ for any $(a_i', \rho_i') \in \Sigma_i^{\sigma,h^t}$ and any $i \in N$,*

$$g_i\left(\sigma^a\left(h^t\right)\right) + \frac{\delta}{1-\delta}\sum_{m\in M} w_i^{\sigma}\left(h^t, m\right)\tilde{p}\left(m|\sigma\left(h^t\right)\right) - \eta \geq$$
$$g_i\left(a_i', \sigma_{-i}^a\left(h^t\right)\right) + \frac{\delta}{1-\delta}\sum_{m\in M} w_i^{\sigma}(h^t, m)\tilde{p}\left(m|(a_i', \rho_i'), \sigma_{-i}\left(h^t\right)\right),$$

This condition means that player i would lose at least η at any public history if she makes any nontrivial deviation.[5,6]

This definition just checks the one-shot deviation constraints at each public history, but all the incentive constraints for the continuation game after each public history are satisfied, because the one-shot deviation principle holds.

Note that η-USPPE σ may assign a suboptimal message off-path, i.e. $\sigma_i^m(h^t, a_i')$ may not be an optimal message strategy when $a_i' \neq \sigma_i^a(h^t)$, since it is just a Nash equilibrium. But we can replace them with an optimal message to obtain a sequential equilibrium that is realization equivalent to σ, because

[5]The incentive constraints for trivial deviations are satisfied by definition.

[6]Another possible formulation of uniformly strict equilibrium would be to require such η-strict incentive uniformly across all the information sets, including the interim stages after observing a private signal.

other players never learn about player i's deviation to a_i' due to the full support assumption.[7]

As an example of η-USPPE, consider any stage game with an η-strict Nash equilibrium. Then repeating this η-strict Nash equilibrium and sending some message independent of histories is an η-uniformly strict PPE.[8] In the following, let $E^\eta(\delta) \subset \mathbb{R}^n$ denote the set of all η-USPPE payoff profiles given δ. In general, η-USPPE may not exist, hence $E^\eta(\delta)$ may be an empty set. The equilibrium payoff set for the standard PPE is compact, but the compactness of $E^\eta(\delta)$ may not be entirely obvious because the set of nontrivial deviations at each public history depends on the continuation payoff profile. However, we can show that $E^\eta(\delta)$ is compact.

Lemma 2 *$E^\eta(\delta)$ is compact.*

Proof. See the Appendix. ■

3 Characterization of Limit Equilibrium Payoff Set

3.1 Constructing The Bounding Set for Equilibrium Payoffs

We characterize the limit η-USPPE payoff set in two steps. In this subsection, we construct a compact set Q^η with the property that $E^\eta(\delta) \subseteq Q^\eta$ for all $\delta \in (0,1)$. In the next subsection, we show that, if $intQ^\eta \neq \emptyset$, then for any $\epsilon > 0$, there exists a nonempty, compact, convex set $W \subseteq Q^\eta$ and $\underline{\delta} \in (0,1)$ such that $W \subseteq E^\eta(\delta)$ for any $\delta \in (\underline{\delta}, 1)$ and the Hausdorff distance between W and Q^η is less than ϵ.

Let $\Lambda = \{\lambda \in \mathbb{R}^n | \, \|\lambda\| = 1\}$ and $e^i = (0,0,..,1,...,0)^\top \in \Lambda$ with the ith coordinate equal to 1. Following the approach of Fudenberg and Levine [7], for each $\lambda \in \Lambda$, we consider the following programming problem $(P^{\lambda,\eta})$.

[7]Note that we do not require any strict incentive for such off-path message strategies and any trivial deviation from the on-path messages, as they do not affect any player's incentive or payoff at all.

[8]Note that any deviation in message after playing the Nash equilibrium is a trivial deviation for this strategy profile.

$$
\begin{aligned}
(P^{\lambda,\eta}) \quad & \sup_{v\in\mathbb{R}^n,\ a\in A,\ \rho\in R,\ x:M\to\mathbb{R}^n} \lambda\cdot v \ s.t. \\
& v = g(a) + E[x(\cdot)\,|(a,\rho)] \\
& g_i(a) + E[x_i(\cdot)\,|(a,\rho)] - \eta \geq g_i(a_i', a_{-i}) + E[x_i(\cdot)\,|(a_i',\rho_i'),(a_{-i},\rho_{-i})] \\
& \forall (a_i',\rho_i') \in \hat{\Sigma}_i^{(a,\rho),x}\ \forall i \in N \\
& \sum_i \lambda_i x_i(m) \leq 0\ \forall m \in M
\end{aligned}
$$

where $\hat{\Sigma}_i^{(a,\rho),x}$ is the set of $(a_i',\rho_i') \in A_i \times R_i$ such that $a_i' \neq a_i$ or player i's unilateral deviation from (a_i,ρ_i) to (a_i',ρ_i') does not change the distribution of $x(\cdot)$ given (a_{-i},ρ_{-i}). Naturally, we call such a deviation nontrivial deviation with respect to x for $(P^{\lambda,\eta})$.

Since the value of the problem is bounded above by $\max_{a\in A}\lambda\cdot g(a)$, it is either some finite value or $-\infty$ when there is no feasible solution.

This programming problem is different from FL's problem in [7] in two aspects because of our uniform strictness requirement. First, an η -wedge is added to the incentive constraints for nontrivial deviations (with respect to $x(\cdot)$). Secondly, we restrict attention to pure actions because uniformly strict equilibrium must be in pure strategies by definition. Note that this problem is independent of δ like FL's problem.

Let $k^\eta(\lambda)$ denote the vale of the supremum for $(P^{\lambda,\eta})$. Let $H^\eta(\lambda) = \{x\in\mathbb{R}^n | \lambda\cdot x \leq k^\eta(\lambda)\}$ be the half space below the hyperplane $\lambda\cdot x = k^\eta(\lambda)$ if $k^\eta(\lambda)$ is finite. $H^\eta(\lambda)$ is an empty set if $k^\eta(\lambda) = -\infty$. Let $Q^\eta = \bigcap_{\lambda\in\Lambda} H^\eta(\lambda)$. The next theorem shows that Q^η is a bound of the equilibrium payoff set given any η and δ.

Theorem 3 *For any $\eta \geq 0$ and any $\delta \in (0,1)$, $E^\eta(\delta) \subseteq Q^\eta$.*

Proof. If $E^\eta(\delta) = \varnothing$, then $E^\eta(\delta) \subseteq Q^\eta$ is trivially true. So suppose that $E^\eta(\delta) \neq \varnothing$ and recall that $E^\eta(\delta)$ is a nonempty compact set by Lemma 2. Fix any $\eta \geq 0$ and pick any $\lambda \in \Lambda$. Let v^* be the η-uniformly strict PPE payoff profile that solves $\max_{v\in E^\eta(\delta)}\lambda\cdot v$. Let σ^* be the equilibrium strategy profile to achieve v^* and $(a^*,\rho^*) \in A\times R$ be the equilibrium action profile and the message strategy profile in the first period. Since σ^* is an η-USPPE, it must satisfy the following conditions:

$$g_i\left(a^*\right)+\frac{\delta}{1-\delta}\sum_{m\in M}w_i^{\sigma^*}\left(m\right)\tilde{p}\left(m|(a^*,\rho^*)\right)-\eta\geq$$
$$g_i\left(a_i',a_{-i}^*\right)+\frac{\delta}{1-\delta}\sum_{m\in M}w_i^{\sigma^*}(m)\tilde{p}(m|(a_i',\rho_i'),(a_{-i}^*,\rho_{-i}^*))\ \forall(a_i',\rho_i')\in\Sigma_i^{\sigma^*,h^1}.$$

Define $x_i^*(m)=\frac{\delta}{1-\delta}\left(w_i^{\sigma^*}(m)-v_i^*\right)$. Then $\sum_m\lambda_i x_i^*(m)\leq 0$ because $w_i^{\sigma^*}(m)\in E^\eta(\delta)$. Since x_i^* is a translation of $\frac{\delta}{1-\delta}w_i^{\sigma^*}$ by a constant, $(v^*,(a^*,\rho^*),x^*)$ satisfies all the η-strict incentive compatibility constraints with respect to the set of nontrivial deviations $\hat{\Sigma}_i^{\sigma^*(h^1),x^*}$ for player i in the programming problem $(P^{\lambda,\eta})$. Finally,

$$\begin{aligned}&g\left(a^*\right)+E[x^*(\cdot)|(a^*,\rho^*)]\\&=g\left(a^*\right)+\frac{\delta}{1-\delta}\sum_{m\in M}\left(w^{\sigma^*}(m)-v^*\right)\tilde{p}\left(m|(a^*,\rho^*)\right)\\&=v^*\end{aligned}$$

and it follows that $(v^*,(a^*,\rho^*),x^*)$ is feasible for $(P^{\lambda,\eta})$ and $k^\eta(\lambda)$ is finite. This implies $k^\eta(\lambda)\geq\lambda\cdot v^*$, hence $E^\eta(\delta)$ is contained in the halfspace $H^\eta(\lambda)$. Since this is true for all $\lambda\in\Lambda$, we have $E^\eta(\delta)\subset\bigcap_{\lambda\in\Lambda}H^\eta(\lambda)=Q^\eta$ for any $\delta\in(0,1)$. ∎

The following lemma, which corresponds to Lemma 3.2. in FL, is useful to assess the possibility of a uniformly strict folk theorem.

Lemma 4 *$k^\eta(-e^i)$ is bounded above by $-\underline{v}_i^\eta$, where*

$$\underline{v}_i^\eta=\min_{a\in A}\left[\max\left\{g_i(a),\max_{a_i'\neq a_i}g_i(a)+\eta\right\}\right].$$

Proof. Suppose that (v,a,ρ,x) is feasible for problem $(P^{-e^i,\eta})$. The last constraint of the problem becomes $x_i(m)\geq 0\ \forall m\in M$. Then player i's payoff

$$v_i=g_i(a)+\sum_{m\in M}x_i(m)\tilde{p}(m|(a,\rho))$$

is bounded from below by $g_i(a)$. By the η-strict incentive constraint, v_i is also bounded from below by $\max_{a_i'\neq a_i}g_i(a_i',a_{-i})+\eta$. Hence v_i is bounded below by

$$\underline{v}_i^\eta=\min_{a\in A}\left[\max\left\{g_i(a),\max_{a_i'\neq a_i}g_i(a_i',a_{-i})+\eta\right\}\right]$$

Therefore, $k^\eta(-e^i)$ is bounded from above by $-\underline{v}_i^\eta$. ■

This $\underline{v}_i^\eta$ coincides with the minmax payoff 0 when $\eta = 0$, but can be strictly positive when $\eta > 0$. As the next lemma shows, it coincides with the minmax payoff if and only if there exists a minmax action profile for player i where player i plays an η-strictly optimal action.

Lemma 5 *For each $i \in N$, $\underline{v}_i^\eta = 0$ if there exists $\underline{a}^i \in A$ such that $g_i(\underline{a}^i) = \min_{a'_{-i}} \max_{a'_i} g_i(a'_i, a'_{-i})$ and $g_i(\underline{a}^i) - g_i(a'_i, \underline{a}^i_{-i}) \geq \eta$ for any $a'_i \neq \underline{a}^i_i$. Furthermore, $\underline{v}_i^\eta > 0$ if there is no such $\underline{a}^i \in A$.*

Proof. Fix i and choose any $a \in A$. If a_i is a best response to a_{-i}, then

$$g_i(a) = \max_{a'_i} g_i(a'_i, a_{-i}) \geq \min_{a'_{-i}} \max_{a'_i} g_i(a'_i, a'_{-i}) = 0.$$

If not, then

$$\max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) + \eta \geq \min_{a'_{-i}} \max_{a'_i} g_i(a'_i, a'_{-i}) + \eta \geq 0.$$

Hence, $\max\left\{g_i(a), \max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) + \eta\right\}$ is nonnegative for any $a \in A$.

Suppose that $\underline{a}^i \in A$ satisfies the conditions of the Lemma. Then $g_i(\underline{a}^i) = 0$ and $\max_{a'_i \neq \underline{a}^i_i} g_i(a'_i, \underline{a}^i_{-i}) + \eta \leq g_i(\underline{a}^i) = 0$. Hence $\underline{v}_i^\eta = 0$ is achieved at $a = \underline{a}^i$.

Suppose that there is no such $\underline{a}^i \in A$. Then η must be strictly positive since the minmax action profile would satisfy the conditions when $\eta = 0$. Take any $a \in A$. If a_i is a best response to a_{-i}, then $g_i(a) \geq \min_{a'_{-i}} \max_{a'_i} g_i(a'_i, a'_{-i}) = 0$. Hence we have either (1) $g_i(a) > 0$ or (2) $g_i(a) = 0$ but $g_i(a) - \max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) < \eta$, which implies $\max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) + \eta > 0$. Hence, $\max\left\{g_i(a), \max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) + \eta\right\}$ is strictly positive in either case.

When a_i is not a best response to a_{-i}, then

$$\max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) + \eta \geq \min_{a'_{-i}} \max_{a'_i} g_i(a'_i, a'_{-i}) + \eta > 0$$

as $\eta > 0$. Since A is a finite set, $\underline{v}_i^\eta = \min_{a \in A}\left[\max\left\{g_i(a), \max_{a'_i \neq a_i} g_i(a'_i, a_{-i}) + \eta\right\}\right]$ must be strictly positive. ■

If this bound $\underline{v}_i^\eta$ is strictly positive, then $k(-e^i) = -\underline{v}_i^\eta < 0$, hence the minmax payoff can never be approximated by any η-USPPE by Theorem 3. So, it is necessary for a folk theorem that an η-strict incentive is provided by the current payoffs at the minmax point.

Also note that this bound may be achieved by some non-minmax action profile when it is strictly positive. If no minmax action profile for player i is η-strictly optimal for i, then some non-minmax action profile $\hat{a} \in A$ may achieve $\underline{v}_i^\eta > 0$ if $g_i(\hat{a})$ is close to 0 and $\max_{a_i' \neq \hat{a}_i} g_i(a_i', \hat{a}_{-i}) + \eta$ is small as well (any deviation from $\hat{a}$ is very costly for player i).

Similarly, we observe that $k^\eta(e^i)$ may be strictly below $\max_a g_i(a)$ unless it is η-strictly optimal for player i to play the action that achieves this value. Otherwise, an additional incentive needs to be provided for player i through some punishment (as $\lambda = e^i$). This necessarily leads to some inefficiency because punishment occurs with positive probability (Green and Porter [10]). Thus it is necessary for a folk theorem that $g_i(a) - g_i(a_i', a_{-i}) \geq \eta$ holds for every $a_i' \neq a_i$ for some action profile a that solves $\max_a g_i(a)$. If this is not the case, then $k^\eta(e^i)$ must be less than $\max_a g_i(a)$ and may be achieved by some action profile that does not solve $\max_a g_i(a)$.

3.2 Limit Result for Equilibrium Payoff Set

3.2.1 Decomposability and Local Decomposability

Our main theorem claims that Q^η provides the limit η-USPPE payoff set when Q^η has an interior point. We prove it by establishing η-uniformly strict versions of many well-known results in [7] and in [1].

We first observe that a set of payoffs can be supported by η-USPPE if it is self-decomposable with respect to η-strict incentive constraints with respect to nontrivial deviations. Given $\delta \in (0,1)$ and $w : M \to \mathbb{R}^n$, we consider the static game $\Gamma^\delta(G, p, w)$, where player i's strategy set is $A_i \times R_i$ and player i's payoff is $(1-\delta)\, g_i(a) + \delta E[w_i(\cdot)|(a,\rho)]$, where w assigns payoffs for each message profile and the expectation is computed with respect to $\tilde{p}(\cdot|a,\rho)$.

Definition 6 *A pair consisting of an action profile $a \in A$ and a profile of message strategies $\rho \in R$ is* η-enforceable *for $\eta > 0$ with respect to nonempty set $W \subset \mathbb{R}^n$ and $\delta \in (0,1)$ if there exists a function $w : M \to W$ such that, for all $i \in I$,*

$$\begin{aligned}&(1-\delta)\, g_i(a) + \delta E[w_i(\cdot)\,|(a,\rho)] - (1-\delta)\eta \\ &\geq (1-\delta)\, g_i(a_i', a_{-i}) + \delta E[w_i(\cdot)\,|(a_i', \rho_i'), (a_{-i}, \rho_{-i})]\ \forall (a_i', \rho_i') \in \hat{\Sigma}_i^{(a,\rho),w}\end{aligned}$$

where $\hat{\Sigma}_i^{(a,\rho),w}$ is the set of nontrivial deviations from (a,ρ) for player i with respect to w. If $v = (1-\delta)\, g(a) + \delta E[w(\cdot)\,|(a,\rho)]$ for some η-enforceable pair (a,ρ) and $w : M \to W$, then we say that v is *η-decomposable* and that

$((a,\rho),w)$ η-decomposes v with respect to W and δ. Define the set of η-decomposable payoffs with respect to W and δ as follows:

$$B(\delta, W, \eta) := \{v \in \mathbb{R}^n | v \text{ is } \eta\text{-decomposable with respect to } W \text{ and } \delta\}.$$

We say that W is *η-self decomposable* with respect to δ if $W \subset B(\delta, W, \eta)$.

It is easy to see that a "uniformly strict" version of Theorem 1 in Abreu, Pearce, and Stacchetti [1] holds here: if W is η-self decomposable with respect to δ, then every $v \in B(\delta, W, \eta)$ can be supported by some η-USPPE. Since the following lemma follows easily from the result in Abreu, Pearce and Stacchetti, its proof is omitted.

Lemma 7 *If a nonempty set $W \subset \mathbb{R}^n$ is bounded and η-self decomposable with respect to $\delta \in (0,1)$, then $B(\delta, W, \eta) \subset E^\eta(\delta)$.*

For the rest of this subsection, we prove that local η-self decomposability of W implies η-self decomposability of W. In the framework of repeated games with imperfect public monitoring, Fudenberg, Levine, and Maskin ([9]) introduced a notion of local self decomposability that is sufficient for self decomposability. Here we prove the corresponding lemma in our setting. We begin with a lemma that establishes a certain monotonicity property of B. It implies that, if W is η-self decomposable with respect to $\delta \in (0,1)$, then W is η-self decomposable for every $\delta' \in (\delta, 1)$.

Lemma 8 *If $W \subseteq \mathbb{R}^n$ is convex and $C \subseteq B(\delta, W, \eta) \cap W$, then $C \subseteq B(\delta', W, \eta)$ for every $\delta' \in (\delta, 1)$.*

Proof. Suppose that $v \in C$. Since $v \in B(\delta, W, \eta)$, v is η-decomposable with respect to W and δ, hence there exists a pair $((a,\rho), w^\delta)$ that η-decomposes v. For any $\delta' > \delta$, define $w^{\delta'} : M \to W$ as the following convex combination of v and w^δ:

$$w^{\delta'}(m) = \frac{\delta' - \delta}{\delta'(1-\delta)} v + \frac{\delta(1-\delta')}{\delta'(1-\delta)} w^\delta(m).$$

Clearly, $w^{\delta'}(m) \in W$ for each $m \in M$ since W is convex. Furthermore, we can show that, for every $\delta' \in (\delta, 1)$, the pair $((a,\rho), w^{\delta'})$ η-decomposes v with respect to W and δ'. To see this, first note that, for all $\delta' \in (\delta, 1)$, and $i \in N$,

$$\begin{aligned}
& (1-\delta') g_i(a) + \delta' E\left[w_i^{\delta'}(\cdot)|(a,\rho)\right] \\
= \;& (1-\delta') g_i(a) + \frac{\delta(1-\delta')}{1-\delta} E\left[w_i^\delta(\cdot)|(a,\rho)\right] + \frac{\delta'-\delta}{1-\delta} v_i \\
= \;& \frac{1-\delta'}{1-\delta}\left\{(1-\delta) g_i(a) + \delta E\left[w_i^\delta(\cdot)|(a,\rho)\right]\right\} + \frac{\delta'-\delta}{1-\delta} v_i \\
= \;& v_i.
\end{aligned}$$

Next, note that for all $(a_i', \rho_i') \in \Sigma_i^{(a,\rho),w^{\delta'}} = \Sigma_i^{(a,\rho),w^{\delta}}$ and $i \in N$,

$$
\begin{aligned}
& (1-\delta')\, g_i(a) + \delta' E[w_i^{\delta'}(\cdot)|(a,\rho)] - (1-\delta')\eta \\
= \; & \frac{(1-\delta')}{(1-\delta)} \left[(1-\delta) g_i(a) + \delta E\left[w_i^{\delta}(\cdot)|(a,\rho)\right]\right] + \frac{\delta'-\delta}{(1-\delta)} v_i - (1-\delta')\eta \\
\geq \; & \frac{(1-\delta')}{(1-\delta)} \left[(1-\delta)\, g_i(a_i', a_{-i}) + \delta E[w_i^{\delta}(\cdot)\,|(a_i',\rho_i'),(a_{-i},\rho_{-i})] + (1-\delta)\eta\right] + \frac{\delta'-\delta}{(1-\delta)} v_i - (1-\delta')\eta \\
= \; & (1-\delta')\, g_i(a_i', a_{-i}) + \delta' \left[\frac{(1-\delta')\,\delta}{\delta'(1-\delta)} E[w_i^{\delta}(\cdot)\,|(a_i',\rho_i'),(a_{-i},\rho_{-i})] + \frac{\delta'-\delta}{\delta'(1-\delta)} v_i\right] \\
= \; & (1-\delta')\, g_i(a_i', a_{-i}) + \delta' E\left[w_i^{\delta'}(\cdot)|(a_i',\rho_i'),(a_{-i},\rho_{-i})\right].
\end{aligned}
$$

Consequently, $\left((a,\rho), w^{\delta'}\right)$ η-decomposes v with respect to W and δ'. Hence $C \subseteq B(\delta', W, \eta)$ for every $\delta' \in (\delta, 1)$ and this completes the proof. ■

Next we introduce local η-self decomposability and show that local η-self decomposability implies η-self decomposability for sufficiently large discount factors.

Definition 9 *A nonempty set $W \subseteq \mathbb{R}^n$ is* locally *η-self decomposable if, for any $v \in W$, there exists $\delta \in (0,1)$ and an open set U containing v such that $U \cap W \subset B(\delta, W, \eta)$.*

Lemma 10 *If $W \subset \mathbb{R}^n$ is compact, convex, and locally η-self decomposable, then there exists a $\underline{\delta} \in (0,1)$ such that W is η-self decomposable with respect to δ for any $\delta \in (\underline{\delta}, 1)$.*

Proof. Choose $v \in W$. Since W is η-locally self decomposable, there exists $\delta_v \in (0,1)$ and an open ball U_v around v such that

$$U_v \cap W \subseteq B(\delta_v, W, \eta).$$

Since W is compact, there exists a finite sub-collection $\{U_{v_k}\}_{k=1}^K$ that covers W. Define $\underline{\delta} = \max_{k=1,\ldots,K}\{\delta_{v_k}\}$. Then

$$U_{v_k} \cap W \subseteq B(\delta_{v_k}, W, \eta) \subseteq B(\delta, W, \eta)$$

for any $\delta \in (\underline{\delta}, 1)$ by Lemma 8 and the convexity of W. Consequently,

$$W = \cup_{k=1}^K (U_{v_k} \cap W) \subseteq B(\delta, W, \eta).$$

for every $\delta \in (\underline{\delta}, 1)$. ■

3.2.2 Local Decomposability of a Smooth Set in The Bounding Set

We call a nonempty compact and convex set in $\mathbb{R}^n$ *smooth* if there exists a unique supporting hyperplane at every boundary point of the set.

The following lemma shows that, if Q^η has an interior point in $\mathbb{R}^n$, then there exists a smooth, compact and convex set in $intQ^\eta$ that is arbitrarily close to Q^η.

Lemma 11 *Suppose that $Q^\eta \subseteq \mathbb{R}^n$ has an interior point. Then, for every $\varepsilon > 0$, there exists a smooth compact and convex set $W' \subset intQ^\eta$ such that the Hausdorff distance between W' and $intQ^\eta$ is at most ε.*

Proof. Choose any $\varepsilon > 0$. Since bounded sets in Euclidean space are totally bounded, there exists a finite set $Z \subseteq intQ^\eta$ such that, for each $v \in intQ^\eta$, there exists $z \in Z$ such that $\|z - x\| < \varepsilon$. Let $W = coZ$. Then W is nonempty, compact and convex. Since Q^η is convex, it follows that $intQ^\eta$ is convex, hence $W \subseteq intQ^\eta$. For each $v \in intQ^\eta$, there exists $z \in W$ such that $\|z - x\| < \varepsilon$, which implies $\sup_{v \in intQ^\eta}[\min_{z \in W} \|z - v\|] \leq \varepsilon$. On the other hand, $\max_{z \in W}[\min_{v \in intQ^\eta} \|z - v\|]$ is clearly 0 since $W \subseteq intQ^\eta$. Hence the Hausdorff distance between W and $intQ^\eta$ is at most ε.

Next we construct W' from W. Since $W \subseteq intQ^\eta$ is a polyhedron and has only a finite number of vertices, we can find a small enough $\epsilon' > 0$ such that, at every $v \in W$, the closed ball $B_v^{\epsilon'}$ of radius ϵ' around v is in $intQ^\eta$. Define $W' = \bigcup_{v \in W} B_v^{\epsilon'}$. Since $W \subseteq W' \subseteq intQ^\eta$, the Hausdorff distance between W' and $intQ^\eta$ is at most ε.

Next we show that W' is a smooth compact convex set. To show that W' is compact, it suffices to show that W' is closed. Suppose that $w_k \in W'$ for each k and $\{w_k\}$ is convergent with limit w^*. For each k, there exists $v_k \in W$ such that $\|v_k - w_k\| \leq \epsilon'$. Since W is compact, we may assume wlog that $\{v_k\}$ is convergent with limit $v^* \in W$. Consequently, $\|v^* - w^*\| \leq \epsilon'$ implying that $w^* \in W'$.

To show that W' is convex, choose any $x, y \in W'$. Then there exists $x', y' \in W$ such that $\|x' - x\| \leq \epsilon'$ and $\|y' - y\| \leq \epsilon'$ by definition of W'. For any $\alpha \in [0, 1]$, $\alpha x' + (1-\alpha)y'$ is in W and the distance between $\alpha x' + (1-\alpha)y'$ and $\alpha x + (1-\alpha)y$ is less than $\alpha \|x' - x\| + (1-\alpha) \|y' - y\| \leq \epsilon'$. So, $\alpha x + (1-\alpha)y \in B_{\alpha x' + (1-\alpha)y'}^{\epsilon'} \subseteq W'$. Therefore W' is convex.

Finally, to see that W' has a unique supporting hyperplane at every boundary point, first note that every boundary point of W' must be a boundary point of $B_v^{\epsilon'}$ for some $v \in W$. Since a supporting hyperplane of a boundary point of W' must be a supporting hyperplane of $B_v^{\epsilon'}$ at the same point and $B_v^{\epsilon'}$ cannot have multiple supporting hyperplanes at any boundary point, the supporting hyperplane must be unique at every boundary point of W'. ■

We now show that any such set W' that approximates Q^η from the inside is η-locally decomposable, which leads to our main result by Lemma 10.

We need two technical lemmas for local decomposability.

Lemma 12 *A smooth, compact and convex set $C \subseteq \mathbb{R}^n$ has non-empty interior in $\mathbb{R}^n$.*

Proof. Suppose otherwise. Then the affine hull of C has dimension less than n. Let S denote the affine hull and, translating if necessary, we may assume that $0 \in C$ and S is a vector subspace of $\mathbb{R}^n$. Since C is smooth, C is not a singleton set. So we can find $p \neq 0 \in C$. Let $x^* \in \arg\max_{x \in C} p \cdot x$. Then $p \cdot x \leq p \cdot x^*$ for all $x \in C$. Hence the set $\{x \in \mathbb{R}^n | p \cdot x = p \cdot x^*\}$ is a supporting hyperplane for C at x^*. Now choose $q \in S^\perp$. Then $q \cdot x = 0$ for all $x \in C$, so the set $\{x \in \mathbb{R}^n | q \cdot x = 0\}$ is a supporting hyperplane for C at x^*. Clearly $\{x \in \mathbb{R}^n | p \cdot x = p \cdot x^*\}$ and $\{x \in \mathbb{R}^n | q \cdot x = 0\}$ are distinct hyperplanes because the former does not include 0 (since $p \cdot x^* \geq p \cdot p > 0$), while the latter does. ■

Lemma 13 *Let $W \subset \mathbb{R}^n$ be a smooth, compact and convex set and let v be a boundary point of W. Let $\lambda^v \neq 0 \in \mathbb{R}^n$ be a normal to the unique supporting hyperplane of W at v, i.e., $\lambda^v \cdot v \geq \lambda^v \cdot x$ for all $x \in W$. Then, for any point $y \in \mathbb{R}^n$ such that $\lambda^v \cdot v > \lambda^v \cdot y$, there exists $\alpha^* \in (0,1)$ such that $(1-\alpha)\, v + \alpha y \in intW$ for any $\alpha \in (0, \alpha^*)$.*

Conversely, if $\lambda^v \in \mathbb{R}^n$ satisfies $\lambda^v \cdot v \geq \lambda^v \cdot x$ for all $x \in W$ and, for any $y \in \mathbb{R}^n$ such that $\lambda^v \cdot v > \lambda^v \cdot y$, there exists $\alpha^* \in (0,1)$ such that $(1-\alpha)\, v + \alpha y \in intW$ for any $\alpha \in (0, \alpha^*)$, then there is the unique supporting hyperplane of W at v and λ^v is its normal vector.

Proof. Translating if necessary, we may assume that $v = 0$. We argue by contradiction. Suppose that there exists $y \in \mathbb{R}^n$ such that $\lambda^v \cdot y < 0$ but for each $\alpha^* \in (0,1)$ there exists $\alpha \in (0, \alpha^*)$ such that $\alpha y \notin intW$. Then there exists a sequence $\{\alpha_k\}$ such that $0 < \alpha_k < 1$, $\alpha_k \to 0$ and $\alpha_k y \notin intW$ for each k. Since $intW$ is non-empty by the previous lemma and convex, there exists for each k a $q_k \neq 0$ such that

$$\frac{q_k}{||q_k||} \cdot x \leq \frac{q_k}{||q_k||} \cdot (\alpha_k y)$$

for all $x \in intW$ by the separating hyperplane theorem. Let $k \to \infty$ and $\frac{q_k}{||q_k||} \to q$ for some $q \neq 0$, extracting a subsequence if necessary. Then $q \cdot x \leq 0$ for all $x \in intW$. Then $q \cdot x \leq 0$ for all $x \in W$, hence q is a normal vector for a supporting hyperplane of W at $v = 0$. To derive a contradiction, we show that $q \neq \beta\lambda^v$ for all $\beta > 0$. To see this, take any $z \in intW$ and note that $\alpha_k^2 z \in intW$ (since $0 \in W$). Therefore

$$\frac{q_k}{||q_k||} \cdot (\alpha_k z) \leq \frac{q_k}{||q_k||} \cdot y$$

implying that $0 \leq q \cdot y$. If $\beta > 0$, then $(\beta\lambda^v) \cdot y < 0$. Consequently, $q \neq \beta\lambda^v$, which contradicts the smoothness of W.

For the converse, $\{x \in \mathbb{R}^n | \lambda^v \cdot x = 0\}$ is clearly supporting hyperplane of W at $v = 0$. Suppose that there is a different supporting hyperplane $\{x \in \mathbb{R}^n | \lambda' \cdot x = 0\}$ of W at $v = 0$, which satisfies $\lambda' \cdot x \leq 0$ for any $x \in W$. Then there must exist $y' \in \mathbb{R}^n$ such that $\lambda' \cdot y' > 0$ and $\lambda^v \cdot y' < 0$. Then $\alpha y'$ is in W for small enough $\alpha \in (0, 1)$ by assumption, but $\lambda' \cdot (\alpha y') > 0$, which is a contradiction. Hence there is the unique supporting hyperplane of W at $v = 0$ and λ^v is its normal vector. ■

Theorem 14 *Suppose that Q^η has an interior point in $\mathbb{R}^n$. For any $\epsilon > 0$, there exists a smooth, compact and convex set $W \subseteq intQ^\eta$ and $\underline{\delta} \in (0, 1)$ such that $W \subset E^\eta(\delta)$ for any $\delta \in (\underline{\delta}, 1)$ and the Hausdorff distance between W and Q^η is at most ϵ.*

Proof. By Lemma 11, there exists a smooth, compact and convex set W in $intQ^\eta$ such that the Hausdorff distance between W and $\text{int}Q^\eta$ is at most ϵ. Since $cl(intQ^\eta) = Q^\eta$, it follows that the Hausdorff distance between W and Q^η is at most ϵ. By Lemma 10, it suffices to show that W is locally η-self decomposable. Take any boundary point w^* of W. Since W is smooth, there is unique $\lambda^* \in \Lambda$ such that $\lambda^* \cdot w^* = \max_{w \in W} \lambda^* \cdot w$. Since $W \subseteq intQ^\eta$, there exists a feasible solution $(v, (a, \rho), x)$ for the programming problem $\left(P^{\lambda^*, \eta}\right)$ such that $\lambda^* \cdot v > \lambda^* \cdot w^*$. Define $x'(m) = x(m) - (v - w^*)$ for each $m \in M$. Then $(w^*, (a, \rho), x')$ is feasible in the programming problem because the η-strict incentive constraints for nontrivial deviations are not affected, $\lambda^* \cdot x'(m) = \lambda^* \cdot x(m) - \lambda^* \cdot (v - w^*) < 0$ for each $m \in M$, and

$$\begin{aligned} & g(a) + E[x'(\cdot)|(a, \rho)] \\ & = g(a) + \sum_{m \in M} x(m)\tilde{p}(m|(a, \rho)) - v + w^* \\ & = w^* \end{aligned}$$

For each δ and m, define $w^\delta(m)$ by $w^\delta(m) := w^* + \frac{1-\delta}{\delta}x'(m)$. If (a,ρ) is played and w^δ is used as the continuation payoff profile, then, for all $(a_i', \rho_i') \in \hat{\Sigma}_i^{(a,\rho),w^\delta} = \hat{\Sigma}_i^{(a,\rho),x'}$ and $i \in N$,

$$\begin{aligned}
&(1-\delta)\, g_i\,(a) + \delta E[w_i^\delta\,(\cdot)\,|(a,\rho)] - (1-\delta)\eta \\
&= (1-\delta)\left(g_i\,(a) + E\left[x_i'(\cdot)|(a,\rho)\right]\right) + \delta w_i^* - (1-\delta)\eta \\
&\geq (1-\delta)\left(g_i\,(a_i', a_{-i}) + E\left[x_i'(\cdot)|(a_i',\rho_i'),(a_{-i},\rho_{-i})\right] + \eta\right) + \delta w_i^* - (1-\delta)\eta \\
&= (1-\delta)\left(g_i\,(a_i', a_{-i}) + E\left[\frac{\delta}{1-\delta}\left(w_i^\delta(\cdot) - w_i^*\right)|(a_i',\rho_i'),(a_{-i},\rho_{-i})\right]\right) + \delta w_i^* \\
&= (1-\delta)\, g_i\,(a_i', a_{-i}) + \delta E\left[w_i^\delta(\cdot)|(a_i',\rho_i'),(a_{-i},\rho_{-i})\right].
\end{aligned}$$

Also note that $(1-\delta)g(a) + \delta E[w^\delta(\cdot)|(a,\rho)] = w^*$.

Next we show that $w^\delta(m)$ is in $intW$ for every m if δ is large enough. Since W is smooth, λ^* is a normal vector of the unique supporting hyperplane of W at w^*. Choose any $\delta' \in (0,1)$. Since $\lambda^* \cdot x'(m) < 0$, it follows that $\lambda^* \cdot w^{\delta'}(m) < \lambda^* \cdot w^*$ for each $m \in M$. Since $\frac{\delta'}{1-\delta'}(w^{\delta'}(m) - w^*) = \frac{\delta}{1-\delta}(w^\delta(m) - w^*)$ for any $\delta, \delta' \in (0,1)$ by definition, we have

$$w^\delta(m) = \left(1 - \frac{(1-\delta)\delta'}{\delta(1-\delta')}\right) w^* + \frac{(1-\delta)\delta'}{\delta(1-\delta')} w^{\delta'}(m)$$

for $\delta \in (\delta', 1)$. Then, for each m, there exists $\underline{\delta}_m \in (0,1)$ such that $w^\delta(m) \in int(W)$ for any $\delta \in (\underline{\delta}_m, 1)$ by Lemma 13. Let $\underline{\delta} = \max_m \underline{\delta}_m$. Then $((a,\rho), w^\delta)$ η-decomposes w^* with respect to $int(W)$ and δ for any $\delta > \underline{\delta}$.

For each $\xi \in \mathbb{R}^n$, let $f^\delta(\xi) = (1-\delta)g(a) + \delta E[w^\delta(\cdot) + \xi|(a,\rho)]$. Then f^δ is continuous, injective and $f(0) = w^*$. Since $w^\delta(m) \in intW$ for each $m \in M$ and M is finite, there exists an open neighborhood V^δ of 0 such that $w^\delta(m) + \xi \in intW$ for each $m \in M$ and each $\xi \in V^\delta$ and $f^\delta(V^\delta) = U^\delta$ is an open neighborhood of w^*. Since f^δ maps V$^\delta$ homeomorphically onto U^δ, it follows that every point $u \in U^\delta$ can be η-decomposed by $((a,\rho), w^\delta + (f^\delta)^{-1}(u))$ with respect to $int(W)$ for each $\delta > \underline{\delta}$. Therefore, $U^\delta \cap W \subseteq U^\delta \subseteq B\,(\delta, intW, \eta) \subseteq B\,(\delta, W, \eta)$ for any $\delta > \underline{\delta}$. A similar argument applies for any $w^* \in intW$. Hence W is locally η-self decomposable. ■

4 Uniformly Strict Folk Theorem

In this section, we prove a folk theorem with η-uniformly strict PPE by showing that Q^η coincides with $V^*(G)$ under certain conditions. In the following, we use ϕ_i to denote a mixed strategy over $A_i \times R_i$. Let $\alpha_i(\phi_i)$ denote the marginal distribution of ϕ_i on A_i. For each (a_{-i}, ρ_{-i}), let

$$\tilde{p}(\cdot|\phi_i, (a_{-i}, \rho_{-i})) = \sum_{(a'_i, \rho'_i) \in A_i \times R_i} \tilde{p}(\cdot|(a'_i, \rho'_i), (a_{-i}, \rho_{-i}))\phi_i\left((a'_i, \rho'_i)\right)$$

and let $\tilde{p}_{-i}(\cdot|\phi_j, (a_{-j}, \rho_{-j}))$ be the marginal distribution of $\tilde{p}(\cdot|\phi_j, (a_{-j}, \rho_{-j}))$ over M_{-i}.

We need the following four conditions on the private monitoring structure and the payoff functions for our folk theorem. Recall that $A(G) \subseteq A$ is the set of action profiles that generate an extreme point in $V(G)$.

Definition 15 (η-detectability) *For each $a \in A(G)$, there exists $\rho \in R$ that satisfies the following condition: for each $i \in N$, if $\tilde{p}(\cdot|\phi_i, (a_{-i}, \rho_{-i})) = \tilde{p}(\cdot|(a, \rho))$ for some $\phi_i \in \Delta\left((A_i \times R_i) \backslash \{(a_i, \rho_i)\}\right)$, then $g_i(a) - g_i(\alpha_i(\phi_i), a_{-i}) \geq \eta$ holds.*

This condition means that if player i's unilateral deviation to a mixed strategy (with 0 probability on (a_i, ρ_i)) cannot be detected, then she must lose at least η in terms of the stage-game expected payoff

When we approximate the minmax point, we need a slightly stronger detectability condition.

Definition 16 (η^*-detectability with respect to i at $a \in A$) *There exists $\rho \in R$ that satisfies the following condition: for each $j \neq i$, if $\tilde{p}_{-i}(\cdot|\phi_j, (a_{-j}, \rho_{-j})) = \tilde{p}_{-i}(\cdot|(a, \rho))$ for some $\phi_j \in \Delta\left((A_j \times R_j) \backslash \{(a_j, \rho_j)\}\right)$, then $g_j(a) - g_j(\alpha_j(\phi_j), a_{-j}) \geq \eta$ holds.*

This condition means that the above η-detectability condition holds for $j \neq i$ without using player i's message.

The next condition means that, if player i's deviation is not linearly independent from some other player's deviation, then she must lose at least η in terms of the stage-game expected payoff.

Definition 17 (η-identifiability) *For each $a \in A(G)$, there exists $\rho \in R$ that satisfies the following condition: for each pair $i \neq j$, if $\tilde{p}(\cdot|\phi_i, (a_{-i}, \rho_{-i})) - \tilde{p}(\cdot|(a, \rho))$ and $\tilde{p}(\cdot|\phi_j, (a_{-j}, \rho_{-j})) - \tilde{p}(\cdot|(a, \rho))$ are not linearly independent for some $\phi_i \in \Delta\left((A_i \times R_i) \backslash \{(a_i, \rho_i)\}\right)$ and $\phi_j \in \Delta\left((A_j \times R_j) \backslash \{(a_j, \rho_j)\}\right)$, then $\min\{g_i(a) - g_i(\alpha_i(\phi_i), a_{-i}), g_j(a) - g_j(\alpha_j(\phi_j), a_{-j})\} \geq \eta$ holds.*

The last conditions require that, for every player i, there exists the best action profile and the minmax action profile, where player i would lose at least η by deviating to any other pure action. Remember Lemma 4 and our discussion following the lemma; we know that they are necessary for the folk theorem.

Definition 18 (η-best response property) *G satisfies η-best response property for $\left\{\overline{a}^i, \underline{a}^i, i \in N\right\} \subset A$ if the following conditions are satisfied for any $i \in N$:*

1. $g_i(\overline{a}^i) = \max_a g_i(a)$ *and* $g_i(\overline{a}^i) - g_i(a'_i, \overline{a}^i_{-i}) \geq \eta$ *for any* $a'_i \neq \overline{a}^i_i$.
2. $g_i(\underline{a}^i) = \min_{a_{-i}} \max_{a_i} g_i(a_i, a_{-i})$ *and* $g_i(\underline{a}^i) - g_i(a'_i, \underline{a}^i_{-i}) \geq \eta$ *for any* $a'_i \neq \underline{a}^i_i$.

It may be useful to compare these conditions to the similar conditions (A1)-(A3) for Theorem 1 in [12]. (A1) requires 0^*-detectability condition at the minmax action profile. We instead assume η^*-detectability condition at the minmax action profile and the best action profile for each player. Kandori and Matsushima [12] assumes (A2) and (A3) for every action profile in $A(G)$, which is a restriction on the distribution of the private signals of any subset of $n-2$ players, whereas we assume η-detectability and η-identifiability for every $a \in A(G)$, which is a restriction on the joint distribution of all messages. η-detectability and η-identifiability are weaker than (A2) and (A3) when the message space is rich enough in the following sense. For η-identifiability, if (A2) and (A3) are satisfied at a and $M_i = S_i$ for every $i \in N$, then η-identifiability is automatically satisfied with truthful message strategies, since the type of linear dependency that appears in the definition of η-identifiability would never occur given (A2) and (A3). For the same reason, (A2) implies η-detectability for every $a \in A(G)$.

As an example of monitoring structure that satisfies our conditions, consider p that satisfies the individual full rank condition for each player with respect to the other players' signals and the pairwise full rank condition for every pair of players with respect to the private signals of the other $n-2$ players. Then (A2) and (A3) are satisfied. In addition, η^*-detectability is trivially satisfied. Hence all our conditions on the monitoring structure (η^*-detectability and η-detectability & η-identifiability) are satisfied.

Using these conditions, we can state our folk theorem with η-USPPE using as follows.

Theorem 19 *Fix any private monitoring game (G,p). Suppose that $intV^*(G) \neq \emptyset$ and G satisfies η-best response property for $\left\{\overline{a}^i, \underline{a}^i, i \in N\right\} \subset A$. If (G,p) satisfies both η-detectability and η-identifiability with the same $\rho^a \in R$ for each $a \in A(G)$ and satisfies η^*-detectability at $\overline{a}_i$ and $\underline{a}_i$ with respect to i for every $i \in N$, then, for any $\epsilon > 0$, there exists a smooth, compact and convex set $W \subseteq intV^*(G)$ and $\underline{\delta} \in (0,1)$ such that $W \subset E^\eta(\delta)$ for any $\delta \in (\underline{\delta}, 1)$ and the Hausdorff distance between W and $V^*(G)$ is at most ϵ.*

We prove this theorem through a series of lemma. We first observe that η-detectability is equivalent to the existence of a transfer x that guarantees η-strict incentive compatibility.

Lemma 20 *(G,p) satisfies η-detectability if and only if for any $a \in A(G)$, there exists ρ such that there exists $x_i : M \to \mathbb{R}$ for each $i \in N$ that satisfies*

$$g_i(a) + \frac{\delta}{1-\delta}\sum_{m\in M} x_i(m)\,\tilde{p}(m|(a,\rho)) - \eta \geq$$
$$g_i(a_i', a_{-i}) + \frac{\delta}{1-\delta}\sum_{m\in M} x_i(m)\tilde{p}(m|(a_i', \rho_i'), (a_{-i}, \rho_{-i}))\ \forall (a_i', \rho_i') \neq (a_i, \rho_i)$$

Since the proof for this result is standard, it is omitted.[9] By the same argument, we can show that η^*-detectability with respect to i is equivalent to the existence of a transfer that does not depend on player i's message and guarantees η-strict incentive for every player other than i.

Lemma 21 *(G,p) satisfies η^*-detectability for i with respect to $a \in A$ if and only if there exists a ρ and, for each $j \neq i$, a function $x_j : M_{-i} \to R$ satisfying*

$$g_j(a) + \frac{\delta}{1-\delta}\sum_{m_{-i}\in M_{-i}} x_j(m_{-i})\,\tilde{p}_{-i}(m_{-i}|(a,\rho)) - \eta \geq$$
$$g_j\left(a_j', a_{-j}\right) + \frac{\delta}{1-\delta}\sum_{m_{-i}\in M_{-i}} x_j(m_{-i})\tilde{p}_{-i}(m_{-i}|(a_j', \rho_j'), (a_{-j}, \rho_{-j}))\ \forall (a_j', \rho_j') \neq (a_j, \rho_j)$$

The next lemma shows that $k^\eta(\lambda)$ is equal to $\max_{a\in A} \lambda \cdot g(a)$ for any regular λ (with at least two nonzero elements) when η-detectability and η-identifiability are satisfied.

Lemma 22 *Suppose that (G,p) satisfies η-detectability and η-identifiability with the same $\rho \in R$ for each $a \in A(G)$. Then, for any $\lambda \in \Lambda \notin \{\pm e^i, i \in N\}$, $k^\eta(\lambda) = \max_{a\in A} \sum_i \lambda_i g_i(a)$.*

[9]For example, see the proof of the the corresponding result in [12] (p. 650).

Proof. Pick any $a^\lambda \in A(G)$ that solves $\max_{a\in A} \sum_i \lambda_i g_i(a)$. By assumption, there is the same $\rho^\lambda \in R$ for which the conditions for η-detectability and η-identifiability are satisfied at a^λ. We show that there exists $x : M \to \mathbb{R}^n$ satisfying the following conditions:

$$\sum_{m\in M} x_i(m)\left(\tilde{p}(m|(a^\lambda,\rho^\lambda)) - \tilde{p}(m|(a'_i,\rho'_i),(a^\lambda_{-i},\rho^\lambda_{-i}))\right) \geq g_i(\alpha'_i, a^\lambda_{-i}) - g_i(a^\lambda) + \eta$$
$$\forall (a'_i,\rho'_i) \neq (a^\lambda_i,\rho^\lambda_i)\ \forall i \in N$$
$$\sum_i \lambda_i x_i(m) = 0\ \forall m \in M.$$

This implies $k^\eta(\lambda) = \sum_i \lambda_i g_i(a^\lambda)$ because $((a^\lambda,\rho^\lambda),x)$ is feasible for the problem $(P^{\lambda,\eta})$ and achieves the upper bound $\sum_i \lambda_i g_i(a^\lambda)$, hence is clearly a maximum point for $(P^{\lambda,\eta})$. Note that every on-path deviation is a nontrivial deviation with respect to transfer x we find. The existence of such $(x_i)_{i\in N}$ is equivalent to the feasibility of the following linear programming problem (with value 0):

$$\min_x 0$$
$$\sum_{m\in M} x_i(m)\left(\tilde{p}(m|(a^\lambda,\rho^\lambda)) - \tilde{p}((a'_i,\rho'_i),(a^\lambda_{-i},\rho^\lambda_{-i}))\right) \geq g_i(a'_i, a^\lambda_{-i}) - g_i(a^\lambda) + \eta$$
$$\forall (a'_i,\rho'_i) \neq (a^\lambda_i,\rho^\lambda_i)\ \forall i \in N$$
$$\sum_{i\in N} \lambda_i x_i(m) = 0\ \forall m \in M$$

The dual problem of this problem is:

$$\max_{q\geq 0,d} \sum_{i\in N} \sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} \left(g_i(a'_i,a^\lambda_{-i}) - g_i(a^\lambda) + \eta\right) q_i((a'_i,\rho'_i))$$
$$\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} \left(\tilde{p}(m|(a^\lambda,\rho^\lambda)) - \tilde{p}(m|(a'_i,\rho'_i),(a^\lambda_{-i},\rho^\lambda_{-i}))\right) q_i((a'_i,\rho'_i)) = \lambda_i d(m)\ \forall m \in M,\ \forall i \in N$$

where $q_i((a'_i,\rho'_i)) \geq 0$ is the multiplier for the strict incentive constraint for $(a'_i,\rho'_i) \neq (a^\lambda_i,\rho^\lambda_i)$ and $d(m) \in \mathbb{R}$ is the multiplier for the λ-"budget balancing" condition for $m \in M$.

By the strong duality theorem, the value of the primal problem is 0 if and only if the value of the dual problem is 0. Take any (q,d) that is feasible for the dual problem. For each i, we consider two cases. First suppose $\lambda_i = 0$. Then the following holds for all $m \in M$:

$$\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} \left(\tilde{p}(m|(a'_i,\rho'_i),(a^\lambda_{-i},\rho^\lambda_{-i})) - \tilde{p}(m|(a^\lambda,\rho^\lambda))\right) q_i((a'_i,\rho'_i)) = 0$$

If $q_i((a'_i,\rho'_i)) = 0$ for each $(a'_i,\rho'_i) \neq (a^\lambda_i,\rho^\lambda_i)$, then the ith term of the objective function is 0. If $q_i((a'_i,\rho'_i)) \neq 0$ for some $(a'_i,\rho'_i) \neq (a^\lambda_i,\rho^\lambda_i)$, then this condition is equivalent to $\tilde{p}(m|\phi'_i,(a^\lambda_{-i},\rho^\lambda_{-i})) = \tilde{p}(m|(a^\lambda,\rho^\lambda))\ \forall m \in M$, where $\phi'_i \in \Delta((A_i \times R_i)\backslash\{(a^\lambda_i,\rho^\lambda_i)\})$ is defined by $\phi'_i((a'_i,\rho'_i)) = \frac{q_i((a'_i,\rho'_i))}{\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} q_i((a'_i,\rho'_i))}$. Note that the ith term of the objective function can be written as

$$\left(\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} q_i((a'_i,\rho'_i))\right) \sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} \left(g_i(\alpha_i(\phi'_i),a^\lambda_{-i}) - g_i(a^\lambda) + \eta\right)$$

which is bounded above by 0 by η-detectability.

Next suppose that $\lambda_i \neq 0$. Then there exists j such that $\lambda_j \neq 0$ since $\lambda \notin \{\pm e^i, i \in N\}$. Consequently, for all $m \in M$ we have

$$\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} \left(\tilde{p}(m|(a^\lambda,\rho^\lambda)) - \tilde{p}(m|(a'_i,\rho'_i),(a^\lambda_{-i},\rho^\lambda_{-i}))\right) q_i((a'_i,\rho'_i)) = \lambda_i d(m)$$
$$\sum_{(a'_j,\rho'_j)\neq(a^\lambda_j,\rho^\lambda_j)} \left(\tilde{p}(m|(a^\lambda,\rho^\lambda)) - \tilde{p}(m|(a'_j,\rho'_j),(a^\lambda_{-j},\rho^\lambda_{-j}))\right) q_j((a'_j,\rho'_j)) = \lambda_j d(m)$$

If $d(m) = 0$ for all m, then we can apply the same argument as before to show that the ith and jth terms of the objective function are at most 0. If $d(m) \neq 0$, then q_i is not identically 0 nor is q_j identically 0. So we can "cross multiply" the two equalities, cancel $d(m)$ and conclude that, for all $m \in M$,

$$\left[\frac{\lambda_j \left(\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} q_i((a'_i,\rho'_i))\right)}{\lambda_i \left(\sum_{(a'_j,\rho'_j)\neq(a^\lambda_j,\rho^\lambda_j)} q_j((a'_j,\rho'_j))\right)}\right] \left(\tilde{p}(m|\phi'_i,(a^\lambda_{-i},\rho^\lambda_{-i})) - \tilde{p}(m|(a^\lambda,\rho^\lambda))\right)$$
$$= \left(\tilde{p}(m|\phi'_j,(a^\lambda_{-j},\rho^\lambda_{-j})) - \tilde{p}(m|(a^\lambda,\rho^\lambda))\right)$$

where ϕ'_i and ϕ'_j are defined by $\phi'_i((a'_i,\rho'_i)) = \frac{q_i((a'_i,\rho'_i))}{\sum_{(a'_i,\rho'_i)\neq(a^\lambda_i,\rho^\lambda_i)} q_i((a'_i,\rho'_i))}$ and $\phi'_j((a'_j,\rho'_j)) = \frac{q_j((a'_j,\rho'_j))}{\sum_{(a'_j,\rho'_j)\neq(a^\lambda_j,\rho^\lambda_j)} q_j((a'_j,\rho'_j))}$ respectively.

Since $\tilde{p}(\cdot|\phi'_i,(a^\lambda_{-i},\rho^\lambda_{-i})) - \tilde{p}(\cdot|(a^\lambda,\rho^\lambda))$ and $\tilde{p}(\cdot|\phi'_j,(a^\lambda_{-j},\rho^\lambda_{-j})) - \tilde{p}(\cdot|(a^\lambda,\rho^\lambda))$ are not linearly independent, it follows from η-identifiability that both

$\sum_{(a_i',\rho_i')\neq(a_i^\lambda,\rho_i^\lambda)}\left(g_i(\alpha_i(\phi_i'),a_{-i}^\lambda)-g_i(a^\lambda)+\eta\right)$ and $\sum_{(a_j',\rho_j')\neq(a_j^\lambda,\rho_j^\lambda)}\left(g_j(\alpha_j(\phi_j'),a_{-j}^\lambda)-g_j(a^\lambda)+\eta\right)$ are bounded above by 0. This implies that the ith term (and the jth term) of the objective function are bounded above by 0.

Hence the ith term of the objective function is bounded above by 0 in either case for any feasible (q,d), implying that the value of the dual problem is bounded above by 0 for any feasible (q,d). Since 0 can be achieved by $q(\cdot)=0$ and $d(\cdot)=0$, the value of the dual problem is exactly 0 as we wanted to show. ■

The next lemma shows that η-best response property and η^* -detectability with respect to i is sufficient to guarantee $k^\eta(e^i)=\max_a g_i(a)$ and $k^\eta(-e^i)=0$.

Lemma 23 *Suppose that G satisfies η-best response property for $\left\{\overline{a}^i,\underline{a}^i,i\in N\right\}\subset A$. Then the following holds for each $i\in N$.*

- *If (G,p) satisfies η^*-detectability with respect to i at $\overline{a}^i$, then $k^\eta(e^i)=\max_a g_i(a)$.*
- *If (G,p) satisfies η^*-detectability with respect to i at $\underline{a}^i$, then $k^\eta(-e^i)=-\min_{a_{-i}}\max_{a_i} g_i(a)=0$.*

Proof. For $\lambda=e^i$, we can find $\overline{a}^i\in A$ such that $g_i(\overline{a}^i)=\max_a g_i(a)$ and $g_i(\overline{a}^i)-g_i(a_i',\overline{a}_{-i}^i)\geq\eta$ for any $a_i'\neq\overline{a}_i^i$ by assumption. Let $\overline{\rho}^i\in R$ be the profile of message strategies for which the conditions for η^*-detectability with respect to i are satisfied at $\overline{a}^i$ for any $j\neq i$. By Lemma 21, for each $j\neq i$, there exists $x_j:M_{-i}\to\mathbb{R}$ such that all the η-strict incentive compatibility conditions are satisfied for any $(a_j',\rho_j')\neq(\overline{a}_j^i,\overline{\rho}_j^i)$. For player i, set $x_i(m)=0$ for all $m\in M$. Then the η-strict incentive compatibility conditions for player i are satisfied for every nontrivial deviation $(a_i',\rho_i')\in\hat{\Sigma}_i^{(\overline{a}^i,\overline{\rho}^i),x}$, since i's message does not affect the transfer x for any player (so deviating in message after the equilibrium action is a trivial deviation). Since $\sum_i\lambda_i x_i(m)=0$ for each m by construction, $(\overline{a}^i,\overline{\rho}^i,x)$ generates an objective function value of $\max_a g_i(a)$ for the problem $(P^{e^i,\eta})$. Clearly this is the largest possible value for $(P^{e^i,\eta})$, hence $k^\eta(e^i)=\max_a g_i(a)$.

For $\lambda=-e^i$, we can find $\underline{a}^i\in A$ such that $g_i(\underline{a}^i)=\min_{a_{-i}}\max_{a_i} g_i(a)=0$ and $g_i(\underline{a}^i)-g_i(a_i',\underline{a}_{-i}^i)\geq\eta$ for any $a_i'\neq\underline{a}_i^i$. Let $\underline{\rho}^i\in R$ be any profile of message strategies for which the conditions for η^*-detectability with respect to i are satisfied at $\underline{a}^i$ for any $j\neq i$. As in the previous case, we can find $x_j:M_{-i}\to\mathbb{R}$ for each $j\neq i$ such that all the η-strict incentive compatibility conditions are satisfied for j. Set $x_i(m)=0$ for all m for player i. Since

$\sum_i \lambda_i x_i(m) = 0$, $(\underline{a}^i, \underline{\rho}^i, x)$ generates an objective function value of $-g_i(\underline{a}^i) = -\min_{a_{-i}} \max_{a_i} g_i(a) = 0$ for $(P^{-e^i,\eta})$.

Since $k(-e^i)$ is bounded from above by 0 by Lemma 4 and Lemma 5, $\underline{a}^i$ solves $(P^{-e^i,\eta})$. Hence $k^\eta(-e^i) = -g(\underline{a}^i) = -\min_{a_{-i}} \max_{a_i} g_i(a) = 0$. ■

Now we complete the proof of Theorem 19. The last two lemmas prove $k^\eta(\lambda) = \max_a \sum_i \lambda_i g_i(a)$ for any $\lambda \notin \{-e^i, i \in N\}$ and $k^\eta(-e_i) = 0$ for every $i \in N$. Since $V^*(G)$ is a compact and convex set, $V^*(G) = \bigcap_\lambda H^\eta(\lambda) = Q^\eta$. Then the theorem follows from Theorem 14 when $intV^*(G) \neq \emptyset$.

5 Discussion

More Strict Incentive Constraints

For our uniformly strict folk theorem, we require a fixed level of strict incentive compatibility at every public history. In terms of average payoff, the strict incentive $(1-\delta)\eta$ converges to 0 as $\delta \to 1$. This means that the loss from a single deviation becomes negligible relative to the size of the total payoff in the limit. We could instead require η-strict incentive compatibility in terms of average payoff. This means that the loss from a deviation is comparable to a permanent payoff shock, say, losing \$1 in all the future periods. To do this, we would replace η in the definition of η-USPPE (Definition 1) with $\frac{\eta}{1-\delta}$. However, it turns out that the set of η-USPPE in this sense becomes empty for any $\eta > 0$ as $\delta \to 1$.

More generally, we can impose $f(\delta)$-strict incentive constraint in terms of average payoff, where $f(\delta)$ may not converge to 0 or converge to 0 more slowly than $(1-\delta)$ as $\delta \to 1$. We can show that, for any such f, the set of "f-uniformly strict" PPE becomes empty for large enough δ. In this sense, our folk theorem cannot be improved in terms of the order of the strict incentive in the limit.

The reason for this is as follows. The effect of the current stage game payoff vanishes at the rate of $(1-\delta)$ as $\delta \to 1$ in terms of average payoff. So, if we like to provide $f(\delta)$-strict incentive with $f(\delta)$ such that $\lim_{\delta\to 1} \frac{f(\delta)}{1-\delta} \to \infty$, it must come from the variation in continuation payoffs.[10] However, the maximum variation of continuation payoffs for player i must vanish at the same rate of $(1-\delta)$ if her continuation payoff $w_i(m)$ is always at least as large as the equilibrium payoff v from the present period. This is because the

[10]For example, it is easy to see that a repetition of any strict Nash equilibrium in the stage game is not $f(\delta)$-USPPE if δ is large enough.

distance between the expected continuation payoff and the equilibrium payoff is $E[w(\cdot)|(a,\rho)] - v = \frac{1-\delta}{\delta}(v - g(a))$, which shrinks to 0 at the rate of $1-\delta$. Hence, to provide $f(\delta)$-strict incentive, continuation payoff must be strictly less than the equilibrium payoff after some message profile, i.e., there exists $\epsilon > 0$ and $m \in M$ such that $w(m) < v - \epsilon$ for any large δ. However this cannot happen at every public history, hence there is no $f(\delta)$-USPPE with such $f(\delta)$ for any large enough discount factor.

Folk Theorem with Double Limits

We prove our folk theorem by fixing a level of strict incentive $\eta > 0$ and letting $\delta \to 1$. If we instead allow η go to 0 and δ go to 1, then we can prove a folk theorem with weaker conditions. When η is small, we can construct Q^η using the minmax action profiles if (G, p) just satisfies η-detectability instead of η^*- detectability at the minmax action profiles. Since Q^η converges to $V^*(G)$ as $\eta \to 0$ and is the limit η-USPPE payoff set (with full dimensionality), we can prove a version of folk theorem only with η-detectability (for minmax action profiles in addition to $A(G)$) and η-identifiability, where η goes to 0 and δ goes to 1 at the same time.

References

[1] Abreu, D., D. Pearce, and E. Stacchetti, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* **59**:1041-1063, 1990.

[2] Aoyagi, M., "Communication Equilibria in Repeated Games with Imperfect Private Monitoring," *Economic Theory,* **25**:455-475, 2005.

[3] Ben-Porath, E. and M. Kahneman, "Communication in Repeated Games with Private Monitoring," *Journal of Economic Theory,* **70**:281-297, 1996.

[4] Bhaskar, V. and I. Obara, "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory,* **102**:40-69, 2002.

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

[6] Ely, J., and J. Välimaki, "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, **102**:84-105, 2002.

[7] Fudenberg, D. and D. K. Levine, "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory,* **62**: 103-135, 1994.

[8] Fudenberg, D. and D. K. Levine, "The Nash-Threats Folk Theorem with Communication and Approximate Common Knowledge in Two Player Games," *Journal of Economic Theory,* **132**:461-473, 2007.

[9] Fudenberg, D., D. K. Levine, and E. Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica,* **62**:997-1039, 1994.

[10] Green, Edward J., and Robert H. Porter, "Noncooperative Collusion under Imperfect Price Information," *Econometrica,* **52**:87-100, 1984.

[11] Hörner, J. and W. Olszewski, "The Folk Theorem for Games with Private, Almost-Perfect Monitoring," *Econometrica,* **74**:1499-1544, 2006.

[12] Kandori, M. and H. Matsushima, "Private Observation and Collusion," *Econometrica*, **66**:627-652, 1998.

[13] Laclau, M. "A Folk Theorem for Repeated Games Played on a Network," *Games and Economic Behavior*, **76**:711-737, 2012.

[14] Laclau, M. “Communication in Repeated Network games with Imperfect Monitoring,” *Games and Economic Behavior*, **87**:136-160, 2014.

[15] Mailath, G. and S. Morris, “Repeated Games with Almost-Public Monitoring,” *Journal of Economic Theory*, **102**:189-228, 2002.

[16] Mailath, G. and W. Olszewski, “Folk Theorems with Bounded Recall under (Almost) Perfect Monitoring,” *Games and Economic Behavior*, **71**:174-192, 2011.

[17] Matsushima, H. “Repeated Games with Private Monitoring: Two Players,” *Econometrica*, **72**: 823-852 2004.

[18] Obara, I. “Folk Theorem with Communication,” *Journal of Economic Theory*, **144**:120-134, 2009.

[19] Piccione, M. “The Repeated Prisoner’s Dilemma with Imperfect Private Monitoring,” *Journal of Economic Theory*, **102**:70-83, 2002.

[20] Renault, J., and T. Tomala, “Repeated Proximity Games,” *International Journal of Game Theory*, **27**:539-559, 1998.

[21] Sekiguchi, T. “Efficiency in Repeated Prisoner’s Dilemma with Private Monitoring,” *Journal of Economic Theory*, **76**:345-361, 1997

[22] Sugaya, T. “Folk Theorem in Repeated Games with Private Monitoring,” *The Review of Economic Studies*, **89**:2201-2256, 2022

[23] Tomala, T. “Perfect Communication Equilibria in Repeated Games with Imperfect Monitoring,” *Games and Economics Behavior*, **67**:682-694, 2009.

Appendix

Proof of Lemma 2

Proof. This is trivial if $E^{\eta}(\delta)$ is an empty set, so suppose that it is not. First, note that $E^{\eta}(\delta)$ is bounded, so we must show that $E^{\eta}(\delta)$ is closed. Take any $v^* \in cl(E^{\eta}(\delta))$. Choose a sequence $v^k \in E^{\eta}(\delta)$ in $\mathbb{R}^n$ that converges to v^*. For each k, let $(a^k, \rho^k) \in A \times R$ be the strategy profile in the first period and $w^k : M \to \mathbb{R}^n$ be the continuation payoff profile from the second period of the equilibrium strategy that supports v^k. Note that $w^k(m) \in E^{\eta}(\delta)$ for all $m \in M$. Then for each $i \in N$,

$$v^k = (1-\delta)g_i\left(a^k\right) + \delta \sum_{m \in M} w_i^k(m)\, \tilde{p}\left(m|a^k, \rho^k\right)$$

Since $A \times R$ is compact and $E^{\eta}(\delta)$ is bounded, we may, extracting a subsequence if necessary, assume that (a^k, ρ^k) and w^k are convergent with respective limits (a^*, ρ^*) and w^*. Furthermore, we may assume that $(a^k, \rho^k) = (a^*, \rho^*)$ for all sufficiently large k. Then

$$v^* = (1-\delta)g_i(a^*) + \delta \sum_{m \in M} w_i^*(m)\, \tilde{p}(m|a^*, \rho^*)\,.$$

and for all sufficiently large k,

$$\begin{aligned} &g_i(a^*) + \frac{\delta}{1-\delta} \sum_{m \in M} w_i^k(m)\, \tilde{p}(m|a^*, \rho^*) - \eta \geq \\ &g_i\left(a_i', a_{-i}^*\right) + \frac{\delta}{1-\delta} \sum_{m \in M} w_i^k(m) \tilde{p}(m|(a_i', \rho_i'), (a_{-i}^*, \rho_{-i}^*)) \end{aligned}$$

for all $(a_i', \rho_i') \in \hat{\Sigma}_i^{(a^*, \rho^*), w^k}$. If $(a_i', \rho_i') \in \hat{\Sigma}_i^{(a^*, \rho^*), w^*}$, then $(a_i', \rho_i') \in \hat{\Sigma}_i^{(a^*, \rho^*), w^k}$ for all sufficiently large k, hence in the limit

$$\begin{aligned} &g_i(a^*) + \frac{\delta}{1-\delta} \sum_{m \in M} w_i^*(m)\, \tilde{p}(m|a^*, \rho^*) - \eta \geq \\ &g_i\left(a_i', a_{-i}^*\right) + \frac{\delta}{1-\delta} \sum_{m \in M} w_i^*(m) \tilde{p}(m|(a_i', \rho_i'), (a_{-i}^*, \rho_{-i}^*)) \end{aligned}$$

for all $(a_i', \rho_i') \in \hat{\Sigma}_i^{(a^*, \rho^*), w^*}$. Since $w^*(m) \in cl(E^{\eta}(\delta))$ for all $m \in M$, it follows that $v^* \in B(\delta, cl(E^{\eta}(\delta)), \eta)$, therefore $cl(E^{\eta}(\delta)) \subseteq B(\delta, cl(E^{\eta}(\delta)), \eta)$. Since $cl(E^{\eta}(\delta))$ is bounded (in fact compact), by η-self decomposability (Lemma7), we can conclude that $cl(E^{\eta}(\delta)) \subseteq E^{\eta}(\delta)$, i.e., $E^{\eta}(\delta)$ is closed. ■