Individual Learning and Belief-Free Equilibria in Repeated Games*

Yuichi Yamamoto†

November 3, 2010

Abstract

This paper considers a new class of two-player repeated games with private monitoring, where the unobservable state of the world influences the payoff functions and the relationships between the distribution of signals and actions played. We focus on *belief-free ex-post equilibria* (*BFXE*), a subset of sequential equilibria that has a tractable structure and several robustness properties. We characterize the limit set of BFXE payoffs as the discount factor converges to one. In addition, under mild identifiability conditions, the limit equilibrium payoff set is isomorphic to the set of maps from states to belief-free equilibrium payoffs for the corresponding known-state game; that is, there are BFXE in which the payoffs are approximately the same as if players learn the true state and play a belief-free equilibrium for that state. As an application, we show that BFXE can approximate efficiency in some economic examples such as investment games with uncertainty. When the signal distribution is weakly conditionally independent, a larger payoff set can be achieved using a variant of BFXE.

Journal of Economic Literature Classification Numbers: C72, C73.

Keywords: repeated game, private monitoring, incomplete information, belief-free equilibrium, ex-post equilibrium, learning.

*I am grateful to my advisors, Attila Ambrus, Susan Athey, and especially Drew Fudenberg for extensive discussions and advice. I also thank Daisuke Hirata, Tomasz Strzalecki, and seminar participants at Harvard University for helpful comments.

†Department of Economics, Harvard University, Cambridge, MA 02138, USA; yamamot@fas.harvard.edu.

1 Introduction

The theory of repeated games provides a framework to investigate how long-term relationships facilitate cooperation in various environments. Recently, considerable progress has been made in the study of games with *imperfect private monitoring*, where players do not directly observe their opponents' actions but instead observe noisy private signals, as in the "secret price-cutting" game of Stigler (1964). Consider an oligopolistic market where firms sell to industrial buyers and interact repeatedly. Price and volume of transaction in such a market are typically determined by bilateral negotiation between a seller and a buyer, so that both price and firm sales are private information. In this situation, a firm's sales level is a noisy information channel of price of the opponents, as it tends to be low if the opponents (secretly) undercut their price.[1]

Under private monitoring, players do not share any common information, so it is not obvious whether they can coordinate to punish a deviator. Nevertheless past work has shown that a long-term relationship still helps provide incentives to cooperate.[2] All this work assumes that players know the distribution of private signals as a function of the actions played, but this assumption is frequently not appropriate. For example, in a secret price-cutting game, a firm may know that its sales level tends to decrease if the opponents undercut their price, but may not know exactly how likely it is to decrease. Or, in a principal-agent model where the principal forms a subjective evaluation on the agent's performance, the agent may know that high effort makes good evaluation likely, but may not know its exact probability. This paper investigates the effect of such uncertainty.[3]

[1] Harrington and Skrzypacz (2010) report that these properties are common to the recent lysine and vitamin markets. Other examples of private monitoring include relational contracts with subjective evaluations (Levin (2003) and Fuchs (2007)) and international trade agreements in the presence of concealed trade barriers (Park (2010)).

[2] For example, Pareto-efficient outcomes can be approximately achieved in the prisoner's dilemma, when observations are nearly perfect (Sekiguchi (1997), Bhaskar and Obara (2002), Piccione (2002), Ely and Välimäki (2002), Hörner and Olszewski (2006), Yamamoto (2009), and Chen (2010)), nearly public (Mailath and Morris (2002), Mailath and Morris (2006), Hörner and Olszewski (2009), and Mailath and Olszewski (2010)), statistically independent (Matsushima (2004) and Yamamoto (2007)), or even fully noisy and correlated (Fong, Gossner, Hörner and Sannikov (2010), Sugaya (2009), and Sugaya (2010b)). Kandori (2002) and Mailath and Samuelson (2006) are excellent surveys. See also Lehrer (1990) for the case of no discounting, and Fudenberg and Levine (1991) for the study of approximate equilibria with discounting.

[3] Note that a sequential equilibrium with the known signal distribution is not necessarily an

Specifically, we study two-player repeated games in which the state of the world, chosen by Nature at the beginning of play, influences the distribution of private signals and/or the payoff functions of the stage game. Players may have (perfect or imperfect) private information about the true state before play begins. Note that the state can affect the payoff functions directly, and can affect it indirectly through the effect on the distribution of signals. For example, in a price-setting oligopoly, the firms tend to have higher expected payoffs at a given price at states where high sales level is likely, and hence even if the payoff to each sales level is known, uncertainty about the distribution of sales levels yields uncertainty about the expected payoffs of the stage game. To the best of the author's knowledge, this is the first to consider such uncertainty under private monitoring.[4]

Since observations are private information in our model, players' posterior beliefs about the true state need not coincide in later periods. In particular, while each player may learn the true state from her private history in the long run, it may not lead to "common learning" in the sense of Cripps, Ely, Mailath, and Samuelson (2008), i.e., the true state may not necessarily be (approximate) common knowledge among players.[5] As a result, even if each player can learn the true state from private signals in the long run, it is not obvious whether players can coordinate to condition their play on the state. These features complicate the verification of the incentive compatibility of a given strategy profile, and make it difficult to characterize the entire equilibrium set. Instead, this paper looks at a tractable subset of Nash equilibria, called *belief-free ex-post equilibria* or *BFXE*. A strategy profile is a BFXE if its continuation strategy constitutes a Nash equilibrium given any state and given any history. In a BFXE, a player's belief about

equilibrium if we introduce small uncertainty about the signal distribution.

[4]Fudenberg and Yamamoto (2010a) and Fudenberg and Yamamoto (2010b) consider similar uncertainty but assume that signals are public information. Wiseman (2010) considers the case where players receive private signals drawn from an unknown signal distribution as in this paper, but he assumes that players observe actions and (almost) public information about the true state. Also, there is an extensive literature on repeated games with unknown payoff functions and perfectly observed actions, notably Forges (1984), Sorin (1984), Hart (1985), Sorin (1985), Aumann and Maschler (1995), Cripps and Thomas (2003), Gossner and Vieille (2003), Wiseman (2005), Hörner and Lovo (2009), and Hörner, Lovo, and Tomala (2009).

[5]Cripps, Ely, Mailath, and Samuelson (2008) consider the situation where players try to learn the unknown state of the world by observing a sequence of private signals over time, and give a condition under which players commonly learn the state. In their model, players do observe private signals, but do not choose actions. On the other hand we consider strategic players, so that their result does not directly apply to our setting.

the true state is irrelevant to her best reply, and hence we do not need to track the evolution of these beliefs over time. This idea is an extension of ex-post equilibria of static games to dynamic setting.[6] Another important property of BFXE is that a player's best reply does not depend on her belief about the opponent's private history, so that we do not need to compute these beliefs as well. This second property is closely related to the concept of *belief-free equilibria* of Ely, Hörner, and Olszewski (2005), which are effective in the study of repeated games with private monitoring and with no uncertainty. Note that BFXE reduce to belief-free equilibria, if the state space is a singleton so that players know the structure of the game.[7]

As mentioned above, the set of BFXE is only a subset of Nash equilibria, and need not always exist (although we show that BFXE exist when players are patient and some additional conditions are satisfied; see Remark 2). Nevertheless the study of BFXE can be motivated by the following considerations. First, BFXE can often approximate the efficient outcome, as we show in several examples. Second, BFXE are robust to any specification of the initial beliefs, just as for ex-post equilibria. That is, BFXE remain equilibria when players are endowed with arbitrary beliefs which need not arise from a common prior. Third, BFXE are robust to any specification of how players update their beliefs. For example BFXE are still equilibria even if players employ non-Bayesian updating of beliefs, or even if each player may observe unmodeled signals that are correlated with the opponent's past private history and/or the true state. Finally, BFXE have a recursive property, in the sense that any continuation strategy profile of a BFXE is also a BFXE. This property greatly simplifies our analysis, and may make our

[6]Some recent works use this "ex-post equilibrium approach" in different settings of repeated games, such as perfect monitoring and fixed states (Hörner and Lovo (2009) and Hörner, Lovo, and Tomala (2009)), public monitoring and fixed states (Fudenberg and Yamamoto (2010a) and Fudenberg and Yamamoto (2010b)), and changing states with an i.i.d. distribution (Miller (2009)).

[7]The idea of belief-free equilibria is proposed by Piccione (2002) and extended by Ely and Välimäki (2002), Ely, Hörner, and Olszewski (2005), and Yamamoto (2007). Its limit equilibrium payoff set is fully characterized by Ely, Hörner, and Olszewski (2005) and Yamamoto (2009). Olszewski (2007) is an introductory survey. Kandori and Obara (2006) show that belief-free equilibria can achieve better payoffs than perfect public equilibria for games with public monitoring. Kandori (2010) proposes a generalization of belief-free equilibria, called weakly belief-free equilibria. Takahashi (2010) construct a version of belief-free equilibria in repeated random matching games. Bhaskar, Mailath, and Morris (2008) investigate the Harsanyi-purifiability of belief-free equilibria. Sugaya and Takahashi (2010) show that belief-free public equilibria of games with public monitoring are robust to private-monitoring perturbations.

approach a promising direction for future research.

In Section 4, we provide a full characterization of the set of BFXE payoffs in the limit as the discount factor goes to one. Since we consider games with two or more possible states, there is often a "trade-off" between equilibrium payoffs for different states; for example, if a player has conflicting interests at different states, then increasing her equilibrium payoff for some states may necessarily lower her equilibrium payoff for other states. Our characterization result builds on the linear programming (LP) technique of Ely, Hörner, and Olszewski (2005), but we need to take into account the effect of this trade-off, which was not present in their analysis.[8] Specifically, we consider a static LP problem whose objective function is a weighted sum of a player's payoffs at different states, and demonstrate that the limit set of BFXE payoffs is characterized by solving these LP problem for all weighting vectors. In particular the trade-offs between equilibrium payoffs for different states are determined by LP problems for weighting vectors that have non-zero weights on two or more states.

Next, in Section 5, we focus on a class of games where the signal distributions are different for different states, which allows each player to learn the true state from a sequence of private signals in the long run, and investigate how it affects equilibrium payoffs. In our model, players do not share any common information, and hence it is not obvious whether players can coordinate their play state by state even if individual state learning is possible. Nevertheless, we establish that the limit set of BFXE payoffs is isomorphic to the set of maps from states to payoffs that are achieved by belief-free equilibria for the corresponding known-state game. That is, there are BFXE where the payoffs are the same as if both players learn the true state and coordinate to play a belief-free equilibrium for that state. Applying this state-learning theorem, we show that there are BFXE approximating the efficient outcome state by state in some economic examples such as investment games with uncertainty.

To understand the intuition behind the state-learning theorem, it is helpful to consider players' behavior on the equilibrium path. Roughly speaking, our equilibrium strategy is constructed in such a way that (i) player i makes player

[8]Fudenberg and Levine (1994) proposes a linear programming characterization of the equilibrium payoff set in repeated games with public monitoring, and this technique is extended by subsequent papers such as Ely, Hörner, and Olszewski (2005), Fudenberg, Levine, and Takahashi (2007), Fudenberg and Yamamoto (2010a), and Fudenberg and Yamamoto (2010b).

$-i$ indifferent over all possible continuation strategies given any history, and (ii) player i controls player $-i$'s payoffs in such a way that player $-i$'s continuation payoff at state ω is close to the target payoff when player i has learned that the true state is likely to be ω. Property (ii) implies that player i's individual state learning is sufficient for player $-i$'s payoff of the entire game to approximate the target payoff state by state. Thus, if each player can individually learn the true state, then both players' payoffs approximate the target payoffs state by state, as stated in the theorem. Also, this strategy profile is indeed incentive compatible, as property (i) assures that a player's play is optimal after every history.

Finally in Section 6, we consider the special case of games with *independent monitoring*, where players observe statistically independent signals conditional on an action profile and a hidden common shock, and show that a larger payoff set can be attained by a variant of BFXE. For games with the known signal distribution and with independent monitoring, past work (Matsushima (2004), Ely, Hörner, and Olszewski (2005), Yamamoto (2007), and Yamamoto (2010)) combines the idea of review strategies of Radner (1985) with belief-free equilibria, and shows that Pareto-efficient outcomes are often approximated. We find that their basic idea extends to games with unknown states, although their constructive proofs do not directly apply. Specifically, we consider T-period LP problems as extensions of the static LP problems in Section 4, and characterize the limit payoff set of review strategies with the belief-free ex-post property using these LP problems. As an application, we show that there is an ex-post efficient equilibrium in a secret price-cutting game where the signal distribution is unknown and independent. This result allows us to show that cartel is self-enforcing even if firms do not know how profitable the market is.

2 Framework

2.1 Model

Given a finite set X, let $\triangle X$ be the set of probability distributions over X, and let $\mathscr{P}(X)$ be the set of non-empty subsets of X, i.e., $\mathscr{P}(X) = 2^X \setminus \{\emptyset\}$. Given a subset W of $\mathbb{R}^n$, let co$W$ denote the convex hull of W.

We consider two-player infinitely repeated games, and the set of players is denoted by $I = \{1, 2\}$. At the beginning of the game, Nature chooses the state of the world ω from a finite set $\Omega = \{\omega_1, \cdots, \omega_o\}$, and then each player obtains private information about the true state ω. Specifically, the set Θ_i of player i's possible private information is a partition of Ω, and given the true state $\omega \in \Omega$, she observes $\theta_i^\omega \in \Theta_i$, where θ_i^ω denotes $\theta_i \in \Theta_i$ such that $\omega \in \theta_i$. For notational convenience, let $\theta^\omega = (\theta_i^\omega)_{i \in I}$. In this setting, private information θ_i^ω allows player i to narrow down the set of possible states. For example, player i knows the state if $\Theta_i = \{(\omega_1), \cdots, (\omega_o)\}$. On the other hand, player i has no information about the state if $\Theta_i = \{\Omega\}$. Given $\theta_i \in \Theta_i$, player i forms a belief about the true state ω, but as we restrict attention to ex-post equilibria, the specification of such a belief is irrelevant.

Each period, players move simultaneously, and player $i \in I$ chooses an action a_i from a finite set A_i. Let $A \equiv \times_{i \in I} A_i$. Given an action profile $a \in A$, each player observes a private signal. Let Σ_i denote a finite set of player i's signals, and let $\Sigma = \times_{i \in I} \Sigma_i$.[9] Let $\pi^\omega(\cdot|a) \in \triangle\Sigma$ be a probability distribution of a signal profile $\sigma = (\sigma_i)_{i \in I} \in \Sigma$ at state ω given an action profile $a \in A$. Also, let $\pi_i^\omega(\cdot|a)$ denote the marginal distribution of $\sigma_i \in \Sigma_i$ at state ω conditional on $a \in A$, that is, $\pi_i^\omega(\sigma_i|a) = \sum_{\sigma_{-i} \in \Sigma_{-i}} \pi^\omega(\sigma|a)$. Player i's realized payoff is $u_i^\omega(a_i, \sigma_i)$, so that her expected payoff at state ω given an action profile a is $g_i^\omega(a) = \sum_{\sigma_i \in \Sigma_i} \pi_i^\omega(\sigma_i|a) u_i^\omega(a_i, \sigma_i)$. We write $\pi^\omega(\alpha)$ and $g_i^\omega(\alpha)$ for the probability distribution and expected payoff when players play a mixed action profile $\alpha \in \times_{i \in I} \triangle A_i$. Similarly, we write $\pi^\omega(a_i, \alpha_{-i})$ and $g_i^\omega(a_i, \alpha_{-i})$ for the probability distribution and expected payoff when player $-i$ plays a mixed action $\alpha_{-i} \in \triangle A_{-i}$. Let $g^\omega(a)$ denote the vector of expected payoffs at state ω give an action profile a.[10]

In the infinitely repeated game, players have a common discount factor $\delta \in (0, 1)$. Let $(a_i^\tau, \sigma_i^\tau)$ be player i's pure action and signal in period τ, and we denote player i's private history from period one to period period $t \geq 1$ by $h_i^t = (a_i^\tau, \sigma_i^\tau)_{\tau=1}^t$. Let $h_i^0 = \emptyset$, and for each $t \geq 0$, let H_i^t be the set of all h_i^t. Also, we

[9]Here we consider a finite Σ_i just for simplicity; our results extend to the case with a continuum of private signals, as in Ishii (2009).

[10]If there are $\omega \in \Omega$ and $\tilde{\omega} \neq \omega$ such that $\theta_i^\omega = \theta_i^{\tilde{\omega}}$ and $u_i^\omega(a_i, \sigma_i) \neq u_i^{\tilde{\omega}}(a_i, \sigma_i)$ for some $a_i \in A_i$ and $\sigma \in \Sigma$, then it might be natural to assume that player i does not observe the realized value of u_i as the game is played; otherwise players might learn the true state from observing their realized payoffs. Since we consider ex-post equilibria, we do not need to impose such a restriction.

denote a pair of t-period histories by $h^t = (h_1^t, h_2^t)$, and let H^t be the set of all h^t. A strategy for player i is defined to be a mapping $s_i : \Theta_i \times \bigcup_{t=0}^{\infty} H_i^t \to \triangle A_i$. Let S_i be the set of all strategies for player i, and let $S = \times_{i \in I} S_i$.

We define the feasible payoff set for a given state ω to be

$$V(\omega) \equiv \text{co}\{g^{\omega}(a) | a \in A\}$$

As in the standard case of a game with a known state, the feasible set is both the set of feasible average discounted payoffs in the infinite-horizon game when players are sufficiently patient and the set of expected payoffs of the stage game that can be obtained when players use of a public randomizing device to implement a distribution over the action profiles.

Next we define the feasible payoff set for the overall game to be

$$V \equiv \times_{\omega \in \Omega} V(\omega).$$

Thus a vector $v \in V$ specifies payoffs for each player and for each state, i.e., $v = ((v_1^{\omega_1}, v_2^{\omega_1}), \cdots, (v_1^{\omega_o}, v_2^{\omega_o}))$. Note that a given $v \in V$ may be generated using different action distributions in each state ω. If players observe ω at the start of the game and are very patient, then any payoff in V can be obtained by a state-contingent strategy of the infinitely repeated game. Looking ahead, there will be equilibria that approximate payoffs in V if the state is *identified* by the signals, so that players learn it over time.

2.2 Belief-Free Ex-Post Equilibrium

This paper studies a special class of Nash equilibria called *belief-free ex-post equilibria*. For each $i \in I$, $s_i \in S_i$, $\theta_i \in \Theta_i$, $t \geq 0$, and $h_i^t \in H_i^t$, let $s_i|_{(\theta_i, h_i^t)}$ denote the continuation strategy induced by s_i when player i observed θ_i and her past private history was h_i^t. For notational convenience, let $s|_{(\theta, h^t)} = (s_i|_{(\theta_i, h_i^t)})_{i \in I}$.

Definition 1. A strategy profile $s \in S$ is a *belief-free ex-post equilibrium* or *BFXE* if for each $\omega \in \Omega$, $t \geq 0$, and $h^t \in H^t$, the continuation strategy profile $s|_{(\theta^{\omega}, h^t)}$ is a Nash equilibrium in the continuation game after (ω, h^t). That is, for each $\omega \in \Omega$, $i \in I$, $t \geq 0$, and $h^t \in H^t$, the continuation strategy $s_i|_{(\theta_i^{\omega}, h_i^t)}$ is a best reply to $s_{-i}|_{(\theta_{-i}^{\omega}, h_{-i}^t)}$ given that the true state is ω.

In BFXE, a player's best reply does not depend on her belief about the true state or about the opponent's private history. Thus we do not need to compute these beliefs for the verification of incentive compatibility, which considerably simplifies the analysis. BFXE reduces to belief-free equilibria of Ely, Hörner, and Olszewski(2005, hereafter EHO) in known-state games, i.e., $|\Omega| = 1$. Note that repetition of a static ex-post equilibrium is a BFXE. Note also that a BFXE might not exist; for example, if there is no static ex-post equilibrium and the discount factor is enough close to zero, then there is no BFXE.

2.3 BFXE with Public Randomization

EHO show that allowing access to public randomization greatly simplifies the analysis of belief-free equilibria. Here we follow this approach, and study BFXE for games with public randomization. We assume that players observe a public signal y at the beginning of every period. Suppose that public signals are i.i.d. draws from the uniform distribution on $Y = [0,1]$. Let y^t denote a public signal in period t, and with abuse of notation, let $h_i^t = (y^\tau, a_i^\tau, \sigma_i^\tau)_{\tau=1}^t$ denote player i's history up to period t. Likewise, let $h^t = (y^\tau, (a_i^\tau, \sigma_i^\tau)_{i\in I})_{\tau=1}^t$ denote a pair of private and public histories up to period t. Let H_i^t be the set of all h_i^t, and H^t be the set of all h^t.

In this setting, a player's play in period $t+1$ is dependent on her own history up to period t and a public signal at the beginning of period $t+1$. Thus a strategy for player i is defined as a mapping $s_i : \Theta_i \times \bigcup_{t=0}^\infty (H_i^t \times Y) \to \triangle A_i$. Let $s_i|_{(\theta_i, h_i^t, y^{t+1})}$ denote the continuation strategy after $(\theta_i, h_i^t, y^{t+1})$, and let $s|_{(\theta, h^t, y^{t+1})} = (s_i|_{(\theta_i, h_i^t, y^{t+1})})_{i\in I}$. The following is a natural extension of the concept of BFXE to games with public randomization.

Definition 2. A strategy profile $s \in S$ is a *BFXE with public randomization* if for each $\omega \in \Omega$, $t \geq 1$, $h^{t-1} \in H^{t-1}$, and $y^t \in Y$, the continuation strategy profile $s|_{(\theta^\omega, h^{t-1}, y^t)}$ is a Nash equilibrium of the infinitely repeated game with the true state ω.

In what follows, we look at a tractable class of BFXE, called *stationary BFXE*. Note that this is parallel to EHO's method; they focus on stationary belief-free equilibria to derive a powerful characterization of the equilibrium payoff set. To give the definition of stationary BFXE, the following notation is useful. For each $i \in I$ and $\theta_i \in \Theta_i$, let $R_i(\theta_i)$ be a non-empty subset of A_i. As will be explained, this $R_i(\theta_i) \subseteq A_i$ is interpreted as the set of "recommended actions" for player i with type θ_i. Let $R = ((R_i(\theta_i))_{\theta_i \in \Theta_i})_{i \in I}$, that is, R specifies a recommended action set for each pair (i, θ_i) of a player and a type. We call such R a *regime*, and let $\mathscr{R}$ be the set of all possible regimes, i.e.,

$$\mathscr{R} = \times_{i \in I} \times_{\theta_i \in \Theta_i} \mathscr{P}(A_i) = \times_{i \in I} \times_{\theta_i \in \Theta_i} (2^{A_i} \setminus \{\emptyset\}).$$

Let $f : [0,1] \to \mathscr{R}$ be a mapping that selects a regime $R \in \mathscr{R}$ contingently on a public signal $y \in [0,1]$, and let $S_i(f)$ be the set of all player i's strategies such that in each period t, player i chooses her action from the recommended set corresponding to the current regime $R = f(y^t)$. That is, for each measurable function $f : [0,1] \to \mathscr{R}$, $S_i(f)$ denotes the set of all $s_i \in S_i$ such that for each θ_i, $t \geq 1$, h_i^{t-1}, y^t, and a_i, $s_i(\theta_i, h_i^{t-1}, y^t)[a_i] = 0$ if a_i is not an element of the (i, θ_i)-component of $f(y^t)$.

Definition 3. A strategy profile $s \in S$ is a *stationary BFXE with respect to* $p \in \triangle\mathscr{R}$ (or *BFXE with respect to* p in short) if there is a function $f : [0,1] \to \mathscr{R}$ such that (i) $\int_{y \in \{y | f(y) = R\}} dy = p(R)$ for each $R \in \mathscr{R}$, (ii) $s_i \in S_i(f)$ for each $i \in I$, and (iii) $\tilde{s}_i|_{(\theta_i^\omega, h_i^{t-1}, y^t)}$ is a best reply to $s_{-i}|_{(\theta_{-i}^\omega, h_{-i}^{t-1}, y^t)}$ for each $i \in I$, $\tilde{s}_i \in S_i(f)$, $\omega \in \Omega$, $t \geq 1$, h^{t-1}, and y^t.

Clause (i) says that a mapping f selects a regime R with probability $p(R)$ in each period. Clause (ii) says that each player chooses her action from the recommended set. Clause (iii) requires that choosing any recommended action is optimal given any history and given any state ω. It follows from clauses (ii) and (iii) that a stationary BFXE with respect to p is a BFXE with public randomization. On the other hand a BFXE with public randomization may not be a stationary BFXE, as we do not allow a player's best reply (a mapping f) to depend on a calendar time t in the the definition of stationary BFXE. Nevertheless, restricting attention to stationary BFXE is without loss of generality as far as players are patient enough and only equilibrium payoffs are concerned; indeed, in the limit as δ goes to one, the payoff set of BFXE with public randomization is equal to the union of the sets of stationary BFXE payoffs over all $p \in \triangle\mathscr{R}$.[11]

[11] The proof is very similar to the on-line appendix of EHO so that we omit it.

When we consider stationary BFXE with respect to p, without loss of generality we can assume that the set of possible public signals is $\mathscr{R}$ (rather than $[0,1]$) and that public signals $(y^t)_{t=1}^{\infty}$ are i.i.d. draws from the distribution $p \in \triangle\mathscr{R}$. We will maintain this assumption in the rest of this paper for notational convenience.

Given a discount factor $\delta \in (0,1)$, let $E^p(\delta)$ denote the set of BFXE payoffs with respect to $p \in \triangle\mathscr{R}$, i.e., $E^p(\delta)$ is the set of all vectors $v = (v_i^{\omega})_{(i,\omega)\in I\times\Omega}$ such that there is $s \in S$ such that s is a stationary BFXE with respect to p and satisfies

$$(1-\delta)E\left[\sum_{t=1}\delta^{t-1}g_i^{\omega}(a^t)\middle| s,\omega,p\right] = v_i^{\omega}$$

for all $i \in I$ and $\omega \in \Omega$. Note that $v \in E^p(\delta)$ specifies the equilibrium payoff for all players and for all possible states. Also, for each $i \in I$, let $E_i^p(\delta)$ denote the set of player i's BFXE payoffs with respect to $p \in \triangle\mathscr{R}$, i.e., E_i^p is the set of all $v_i = (v_i^{\omega})_{\omega\in\Omega} \in \mathbb{R}^{|\Omega|}$ such that there is a BFXE with respect to p such that player i's equilibrium payoff at state ω is v_i^{ω} for each $\omega \in \Omega$.

The following proposition asserts that given a $p \in \triangle\mathscr{R}$, stationary BFXE with respect to p are interchangeable. To see the reason, let s and $\tilde{s}$ be stationary BFXE with respect to p. By the definition of stationary BFXE, choosing a recommended action in every period is a best reply to $\tilde{s}_{-i}|_{\tilde{h}_{-i}^t}$ for any t and $\tilde{h}_{-i}^t$, and thus playing $s_i|_{h_i^t}$ is a best reply to $\tilde{s}_{-i}|_{\tilde{h}_{-i}^t}$ for any t, h_i^t, and $\tilde{h}_{-i}^t$. Likewise $\tilde{s}_i|_{h_i^t}$ is a best reply to $s_{-i}|_{\tilde{h}_{-i}^t}$ for any t, h_i^t, and $\tilde{h}_{-i}^t$. Therefore both $(s_1,\tilde{s}_2)$ and $(\tilde{s}_1,s_2)$ are stationary BFXE.

Proposition 1. *Let $p \in \triangle\mathscr{R}$, and let s and $\tilde{s}$ be stationary BFXE with respect to p. Then, the profiles $(s_1,\tilde{s}_2)$ and $(\tilde{s}_1,s_2)$ are also stationary BFXE with respect to p.*

The next proposition states that given a p, the equilibrium payoff set has a product structure. This conclusion follows from the fact that stationary BFXE are interchangeable: To see this, fix a p, and let s be a stationary BFXE with payoff $v = (v_1,v_2)$, and $\tilde{s}$ be a stationary BFXE with payoff $\tilde{v} = (\tilde{v}_1,\tilde{v}_2)$. Since stationary BFXE are interchangeable, $(s_1,\tilde{s}_2)$ is also a stationary BFXE, and hence player 1 is indifferent between s_1 and $\tilde{s}_1$ against $\tilde{s}_2$. This implies that player 1's payoff from $(s_1,\tilde{s}_2)$ is equal to $\tilde{v}_1$. Also, player 2 is indifferent between s_2 and $\tilde{s}_2$ against s_1, so that her payoff from $(s_1,\tilde{s}_2)$ is equal to v_2. Therefore $(s_1,\tilde{s}_2)$ is a stationary BFXE with payoff $(\tilde{v}_1, v_2)$. Likewise, $(\tilde{s}_1, s_2)$ is a stationary BFXE with payoff $(v_1, \tilde{v}_2)$. This argument shows that the equilibrium payoff set has a product structure, i.e., if v and $\tilde{v}$ are equilibrium payoffs then $(\tilde{v}_1, v_2)$ and $(v_1, \tilde{v}_2)$ are also equilibrium payoffs.

Proposition 2. *For any $\delta \in (0,1)$ and any $p \in \triangle\mathscr{R}$, $E^p(\delta) = \times_{i\in I} E_i^p(\delta)$.*

Since the equilibrium payoff set $E^p(\delta)$ has a product structure, one may expect that we can characterize the equilibrium payoff set for each player separately. In the next section, we show that this conjecture is true and develop a general method to compute the set of stationary BFXE payoffs.

Remark 1. It may be noteworthy that Propositions 1 and 2 are true only for two-player games. To see this, let s and $\tilde{s}$ be stationary BFXE with respect to p in a three-player game, and consider a profile $(\tilde{s}_1, s_2, s_3)$. As in the two-player case, $\tilde{s}_1$ is a best reply to (s_2, s_3). However, s_2 is not necessarily a best reply to $(\tilde{s}_1, s_3)$, since $\tilde{s}_1$ can give right incentives to player 2 only when player 3 plays $\tilde{s}_3$. Therefore $(\tilde{s}_1, s_2, s_3)$ is not necessarily a BFXE. Since Propositions 1 and 2 are key ingredients in the following sections, it is not obvious whether the theorems in the following sections extend to games with more than two players. A similar problem arises in the study of belief-free equilibria in known-state games; see Yamamoto (2009).

3 Individual Ex-Post Self-Generation

Abreu, Pearce, and Stachetti (1990) gives a precise characterization of the set of prefect public equilibrium payoffs for repeated games with public monitoring, using the recursive property of prefect public equilibria. EHO extend this idea to the case of private monitoring, and provides a recursive characterization of the set of belief-free equilibrium payoffs. In this section, we adapt their technique and develop a general method to compute the set of stationary BFXE payoffs.

By definition, any continuation strategy of a stationary BFXE is also a stationary BFXE. Thus a stationary BFXE specifies BFXE continuation play after any one-period history (y, a, σ). Let $w(y, a, \sigma) = (w_i^\omega(y, a, \sigma))_{(i,\omega)\in I\times\Omega}$ denote the continuation payoffs corresponding to one-period history (y, a, σ). Note that player i's continuation payoff $w_i^\omega(y,a,\sigma)$ at state ω is independent of (a_i,σ_i), as the continuation play is an equilibrium given any (a_i,σ_i); thus we write $w_i^\omega(y,a_{-i},\sigma_{-i})$ for player i's continuation payoff. Let $w_i^\omega(y,a_{-i}) = (w_i^\omega(y,a_{-i},\sigma_{-i}))_{\sigma_{-i}\in\Sigma_{-i}}$, and we write $\pi_{-i}^\omega(a)\cdot w_i^\omega(y,a_{-i})$ for player i's expected continuation payoff at state ω given a public signal y and an action profile a. (Recall that $\pi_{-i}^\omega(a)$ is the marginal distribution of player $-i$'s private signals at state ω.) Also, let $w_i(y,a_{-i},\sigma_{-i}) = (w_i^\omega(y,a_{-i},\sigma_{-i}))_{\omega\in\Omega}$.

Let $\vec{\alpha}_i = (\alpha_i^{R,\theta_i})_{(R,\theta_i)\in\mathscr{R}\times\Theta_i}$ be such that $\alpha_i^{R,\theta_i} \in \triangle R_i(\theta_i)$ for each $R\in\mathscr{R}$ and $\theta_i\in\Theta_i$, and we call it *player i's action plan*. In words, an action plan $\vec{\alpha}_i$ specifies what action to play for each public signal $R\in\mathscr{R}$ and for each type $\theta_i\in\Theta_i$, in such a way that the specified (possibly mixed) action α_i^{R,θ_i} is chosen from the recommended set $\triangle R_i(\theta_i)$. Let $\triangle\vec{A}_i$ denote the set of all such player i's action plans $\vec{\alpha}_i$. That is, $\triangle\vec{A}_i = \times_{R\in\mathscr{R}}\times_{\theta_i\in\Theta_i}\triangle R_i(\theta_i)$.

For a payoff vector $v_i\in\mathbb{R}^{|\Omega|}$ to be a BFXE payoff, it is necessary that v_i is an average of today's payoff and the (expected) continuation payoff, and that player i is willing to choose actions recommended by a public signal y in period one. This motivates the following definition:

Definition 4. For $\delta\in(0,1)$, $W_i\subseteq\mathbb{R}^{|\Omega|}$, and $p\in\triangle\mathscr{R}$, player i's payoff vector $v_i = (v_i^\omega)_{\omega\in\Omega}\in\mathbb{R}^{|\Omega|}$ is *individually ex-post generated with respect to* (δ,W_i,p) if there is player $-i$'s action plan $\vec{\alpha}_{-i}\in\triangle\vec{A}_{-i}$ and a function $w_i:\mathscr{R}\times A_{-i}\times\Sigma_{-i}\to W_i$ such that

$$v_i^\omega = \sum_{R\in\mathscr{R}} p(R)\sum_{a_{-i}\in A_{-i}}\alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i})\hat{w}_i^\omega(\delta,R,a_i^R,a_{-i}) \tag{1}$$

for all $\omega\in\Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ satisfying $a_i^R\in R_i(\theta_i^\omega)$ for each $R\in\mathscr{R}$, and

$$v_i^\omega \geq \sum_{R\in\mathscr{R}} p(R)\sum_{a_{-i}\in A_{-i}}\alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i})\hat{w}_i^\omega(\delta,R,a_i^R,a_{-i}) \tag{2}$$

for all $\omega\in\Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ satisfying $a_i^R\in A_i$ for each $R\in\mathscr{R}$. Here, for each $\omega\in\Omega$, $R\in\mathscr{R}$, and $a\in A$,

$$\hat{w}_i^\omega(\delta,R,a) = (1-\delta)g_i^\omega(a) + \delta\pi_{-i}^\omega(a)\cdot w_i^\omega(R,a_{-i}),$$

that is, $\hat{w}_i^\omega(\delta,R,a)$ denotes player i's total payoff at state ω given that today's public signal and action profile is (R,a) and the future continuation payoff is $w_i^\omega(R,a_{-i},y_{-i})$.

The first constraint is "adding-up" condition, meaning that for each state ω, the target payoff v_i^ω is exactly achieved if player i chooses an action from the recommended set $R_i(\theta_i^\omega) \subseteq A_i$ contingently on a public signal R. The second constraint is ex-post incentive compatibility, which implies that player i has no incentive to deviate from such recommended actions.

For each $\delta \in (0,1)$, $i \in I$, $W_i \subseteq \mathbb{R}^{|\Omega|}$, and $p \in \triangle\mathscr{R}$, let $B_i^p(\delta, W_i)$ denote the set of all player i's payoff vectors $v_i \in \mathbb{R}^{|\Omega|}$ individually ex-post generated with respect to (δ, W_i, p).

Definition 5. A subset W_i of $\mathbb{R}^{|\Omega|}$ is *individually ex-post self-generating with respect to* (δ, p) if $W_i \subseteq B_i^p(\delta, W_i)$.

Using the concept of self-generation, the following two propositions provide a recursive characterization of the set of stationary BFXE payoffs for any discount factor $\delta \in (0,1)$. Proposition 3, which is a counterpart to the second half of Proposition 2 of EHO, asserts that the equilibrium payoff set is a fixed point of the operator B_i^p. Proposition 4 is a counterpart to the first half of Proposition 2 of EHO, and shows that any bounded and individually ex-post self-generating set is a subset of the equilibrium payoff set. Taken together, it turns out that the set of BFXE payoffs is the largest set of individually ex-post self-generating set.

Proposition 3. *For every* $\delta \in (0,1)$ *and* $p \in \triangle\mathscr{R}$, $E^p(\delta) = \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$.

The formal proof is given in Appendix. The result comes from the recursive property of equilibria as in Abreu, Pearce, and Stachetti (1990) and EHO. The inclusion $E^p(\delta) \subseteq \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$ follows from the fact that if a strategy profile is a stationary BFXE then its continuation strategy profile from period two on is also a stationary BFXE. Conversely $E^p(\delta) \supseteq \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$ follows, because a strategy profile is a stationary BFXE if the incentive compatibility constraint in period one is satisfied and its continuation strategy profile from period two on is a stationary BFXE.

Proposition 4. *For each* $i \in I$, *let* W_i *be a subset of* $\mathbb{R}^{|\Omega|}$ *that is bounded and individually ex-post self-generating with respect to* (δ, p). *Then* $\times_{i\in I} W_i \subseteq E^p(\delta)$.

See Appendix for the formal proof. Again the idea is similar to Abreu, Pearce, and Stachetti (1990) and EHO. To prove this proposition, we recursively construct a candidate strategy profile, and then prove that it is indeed a stationary BFXE with the specified equilibrium payoff. The key is that when W_i is individually ex-post self-generating, the continuation payoffs $w_i(y, a_{-i}, \sigma_{-i})$ used to enforce $v_i \in W_i$ have the property that the payoff vector $w_i(y, a_{-i}, \sigma_{-i})$ can in turn be ex-post generated using a single next-period action plan $\vec{\alpha}_{-i}$, which says to play the same action for different states ω and $\tilde{\omega}$ if player $-i$ cannot distinguish these two states by her initial information θ_{-i}.

4 Characterizing the Limit Set of Equilibrium Payoffs

For known-state games (i.e., the case of $|\Omega| = 1$), EHO provide a simple characterization of the set of belief-free equilibrium payoffs in the limit as the discount factor goes to one. Specifically, they show that the maximum and minimum of the limit set of belief-free equilibrium payoffs are computed by solving linear programming problems. However, computing the maximum and minimum of equilibrium payoffs is not sufficient to determine the entire set of equilibrium payoffs for games with two or more possible states, as there can be a trade-off between equilibrium payoffs for different states. For example, increasing a player's equilibrium payoff for some state might necessarily lower her equilibrium payoff for different states, if she has conflicting interests at different states.[12] Taking this problem into account, in this section we demonstrate that the limit set of BFXE payoffs is the intersection of "maximal half-spaces" in various directions. This is an extension of the linear programming characterization of the limit payoff set of PPE, which is proposed by Fudenberg and Levine (1994). (However, note that

[12] Here is a more concrete example. Suppose that there are two states ω_1 and ω_2. In each stage game, player 1 chooses either U or D, and player 2 chooses L or R. After choosing actions, player 1 observes both the true state and the actions played, while player 2 observes only the actions. The stage game payoffs are as follows:

	L	R
U	2, 0	1, 0
D	0, 0	0, 0

	L	R
U	1, 0	2, 0
D	0, 0	0, 0

Note that D is dominated by U at both states, and hence player 1 always chooses U in any BFXE. On the other hand, any strategy profile s where player 1 chooses the pure action U after every history is a BFXE. Therefore, for any δ, player 1's equilibrium payoff set $E_1(\delta)$ is a convex combination of $(1,2)$ and $(2,1)$. So increasing player 1's equilibrium payoff at state ω_1 lowers her equilibrium payoff at ω_2.

the effect of trade-offs between equilibrium payoffs for different states may not appear if each player can learn the true state from private signals; see Section 5.)

4.1 Linear Programming Problem and Bound of $E^p(\delta)$

In this subsection, we provide a bound on the set of BFXE payoffs, by considering a linear programming (LP) problem for each direction λ_i where each component λ_i of the vector λ_i corresponds to the weight attached to player i's payoff at state ω. In particular, trade-offs between equilibrium payoffs for different states are characterized by solving LP problems for "cross-state" directions λ_i that have two or more non-zero components (i.e., directions λ_i that put non-zero weights to two or more states).

Let Λ_i be the set of all $\lambda_i = (\lambda_i^\omega)_{\omega\in\Omega} \in \mathbb{R}^{|\Omega|}$ such that $|\lambda_i| = 1$. For each $R \in \mathscr{R}$, $i \in I$, $\delta \in (0,1)$, $\vec{\alpha}_{-i} \in \triangle\vec{A}_{-i}$, and $\lambda_i \in \Lambda_i$, consider the following LP problem.

$$k_i^p(\vec{\alpha}_{-i},\lambda_i,\delta) = \max_{\substack{v_i\in\mathbb{R}^{|\Omega|} \\ w_i:\mathscr{R}\times A_{-i}\times\Sigma_{-i}\to\mathbb{R}^{|\Omega|}}} \lambda_i\cdot v_i \qquad \text{subject to}$$

(i) (1) holds for all $\omega\in\Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ s.t. $a_i^R \in R_i(\theta_i^\omega)$ for each $R\in\mathscr{R}$,

(ii) (2) holds for all $\omega\in\Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ s.t. $a_i^R \in A_i$ for each $R\in\mathscr{R}$,

(iii) $\lambda_i\cdot v_i \geq \lambda_i\cdot w_i(R,a_{-i},\sigma_{-i})$ for all $R\in\mathscr{R}$, $a_{-i}\in A_{-i}$, and $\sigma_{-i}\in\Sigma_{-i}$.

If there is no (v_i,w_i) satisfying the constraints, let $k_i^p(\vec{\alpha}_{-i},\lambda_i,\delta) = -\infty$. If for every $\overline{k} > 0$ there is (v_i,w_i) satisfying all the constraints and $\lambda_i\cdot v_i > \overline{k}$, then let $k_i^p(\vec{\alpha}_{-i},\lambda_i,\delta) = \infty$. With an abuse of notation, when p is a unit vector such that $p(R) = 1$ for some regime R, we denote the maximal score by $k_i^R(\vec{\alpha}_{-i},\lambda_i)$.

As we have explained in the previous section, (i) is the "adding-up" constraint, and (ii) is ex-post incentive compatibility. Constraint (iii) requires that the continuation payoffs lie in the half-space corresponding to direction λ_i and payoff vector v_i. Thus the solution $k_i^p(\vec{\alpha}_{-i},\lambda_i,\delta)$ to this problem is the maximal score toward direction λ_i that is individually ex-post generated by the half-space corresponding to direction λ_i and payoff vector v_i. Constraint (iii) allows "utility transfer across states" (in other words, "learning and state-contingent play"), which captures the following scenario: If player $-i$ learns from her private signal σ_{-i} that ω is likely to be the true state, then she may choose a continuation strategy that yields high payoffs to player i at state ω, which may necessarily lower player i's payoff at state $\tilde{\omega}$. Note that this issue do not appear in EHO, as they study known-state games.

For each $\omega \in \Omega$, $R \in \mathscr{R}$, a_{-i}, and $\sigma_{-i} \in \Sigma_{-i}$, let

$$x_i^\omega(R, a_{-i}, \sigma_{-i}) = \frac{\delta}{1-\delta}(w_i^\omega(R, a_{-i}, \sigma_{-i}) - v_i^\omega).$$

Also, in order to simplify our notation, let $x_i^\omega(R, a_{-i}) = (x_i^\omega(R, a_{-i}, \sigma_{-i}))_{\sigma_{-i} \in \Sigma_{-i}}$ and let $x_i(R, a_{-i}, \sigma_{-i}) = (x_i^\omega(R, a_{-i}, \sigma_{-i}))_{\omega \in \Omega}$. Arranging constraints (i) through (iii), we can transform the above problem to:

(LP-Individual) $$\max_{\substack{v_i \in \mathbb{R}^{|\Omega|} \\ x_i : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to \mathbb{R}^{|\Omega|}}} \lambda_i \cdot v_i \quad \text{subject to}$$

(i) $$v_i^\omega = \sum_{R \in \mathscr{R}} p(R) \sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R, \theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^\omega(a_i^R, a_{-i}) \\ +\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $\omega \in \Omega$ and $(a_i^R)_{R \in \mathscr{R}}$ s.t. $a_i^R \in R_i(\theta_i^\omega)$ for each $R \in \mathscr{R}$,

(ii) $$v_i^\omega \geq \sum_{R \in \mathscr{R}} p(R) \sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R, \theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^\omega(a_i^R, a_{-i}) \\ +\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $\omega \in \Omega$ and $(a_i^R)_{R \in \mathscr{R}}$ s.t. $a_i^R \in A_i$ for each $R \in \mathscr{R}$,

(iii) $\lambda_i \cdot x_i(R, a_{-i}, \sigma_{-i}) \leq 0$, for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$.

Since δ does not appear in constraints (i) through (iii) of (LP-Individual), the score $k_i^p(\vec{\alpha}_{-i}, \lambda_i, \delta)$ is independent of δ. Thus we will denote it by $k_i^p(\vec{\alpha}_{-i}, \lambda_i)$. Note also that, as in EHO, only the marginal distribution π_{-i} matters in (LP-Individual); that is, the score $k_i^p(\vec{\alpha}_{-i}, \lambda_i)$ depends on the signal distribution π only through the marginal distribution π_{-i}.

Now let

$$k_i^p(\lambda_i) = \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} k_i^p(\vec{\alpha}_{-i}, \lambda)$$

be the highest score that can be approximated in direction λ_i by any choice of $\vec{\alpha}_{-i}$. For each $\lambda_i \in \Lambda_i$ and $k_i \in \mathbb{R}$, let $H_i(\lambda_i, k_i) = \{v_i \in \mathbb{R}^{|\Omega|} | \lambda_i \cdot v_i \leq k_i\}$. Let $H_i(\lambda_i, k_i) = \mathbb{R}^{|\Omega|}$ for $k_i = \infty$, and $H_i(\lambda_i, k_i) = \emptyset$ for $k_i = -\infty$. Then let

$$H_i^p(\lambda_i) = H_i(\lambda_i, k_i^p(\lambda_i))$$

be the maximal half-space in direction λ_i, and let

$$Q_i^p = \bigcap_{\lambda_i \in \Lambda_i} H_i^p(\lambda_i)$$

be the intersection of half-spaces over all λ_i. Let

$$Q^p = \times_{i \in I} Q_i^p.$$

Lemma 1.

- **(a)** $k_i^p(\vec{\alpha}_{-i}, \lambda_i) = \sum_{R \in \mathscr{R}} p(R) k_i^R(\vec{\alpha}_{-i}, \lambda_i)$.
- **(b)** $k_i^p(\lambda_i) = \sum_{R \in \mathscr{R}} p(R) k_i^R(\lambda_i)$.
- **(c)** Q_i^p *is bounded.*

Proof. Inspecting the set of the constraints in the transformed problem, we can check that solving this LP problem is equivalent to find the continuation payoffs $(w_i^\omega(R, a_{-i}, \sigma_{-i}))_{(\omega, a_{-i}, \sigma_{-i})}$ for each regime R in isolation. This proves part (a).

Note that the maximal score $k_i^R(\vec{\alpha}_{-i}, \lambda_i)$ is dependent on an action plan $\vec{\alpha}_{-i}$ only through $(\alpha_{-i}^{R, \theta_{-i}})_{\theta_{-i} \in \Theta_{-i}}$, and the remaining components $(\alpha_{-i}^{\tilde{R}, \theta_{-i}})_{\theta_{-i} \in \Theta_{-i}}$ for $\tilde{R} \neq R$ are irrelevant. Therefore, we have

$$\sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} \sum_{R \in \mathscr{R}} p(R) k_i^R(\vec{\alpha}_{-i}, \lambda_i) = \sum_{R \in \mathscr{R}} p(R) \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} k_i^R(\vec{\alpha}_{-i}, \lambda_i)$$

for any $p \in \triangle \mathscr{R}$. Using this and part (a), we obtain

$$\begin{aligned} k_i^p(\lambda_i) &= \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} k_i^p(\vec{\alpha}_{-i}, \lambda_i) \\ &= \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} \sum_{R \in \mathscr{R}} p(R) k_i^R(\vec{\alpha}_{-i}, \lambda_i) \\ &= \sum_{R \in \mathscr{R}} p(R) \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} k_i^R(\vec{\alpha}_{-i}, \lambda_i) \\ &= \sum_{R \in \mathscr{R}} p(R) k_i^R(\lambda_i) \end{aligned}$$

so that part (b) follows.

To prove part (c), consider $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega \neq 0$ for some $\omega \in \Omega$ and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$. Then from constraint (i) of (LP-Individual),

$$\lambda_i \cdot v_i = \lambda_i^\omega v_i^\omega = \lambda_i^\omega \sum_{R \in \mathscr{R}} p(R) \sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^\omega(a_i^R, a_{-i}) \\ +\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $(a_i^R)_{R \in \mathscr{R}}$ such that $a_i^R \in R_i(\theta_i^\omega)$ for each $R \in \mathscr{R}$. Since constraint (iii) of (LP-Individual) implies that $\lambda_i^\omega \pi_{-i}^\omega(a) \cdot x_i^\omega(R, a_{-i}) \leq 0$ for all $a \in A$ and $R \in \mathscr{R}$, it follows that

$$\lambda_i \cdot v_i \leq \max_{a \in A} \lambda_i^\omega g_i^\omega(a).$$

Thus the maximal score for this λ_i is bounded. Let Λ_i^* be the set of $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega \neq 0$ for some $\omega \in \Omega$ and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$. Then the set $\bigcap_{\lambda_i \in \Lambda_i^*} H_i^p(\lambda_i)$ is bounded. This proves part (c), since $Q_i^p \subseteq \bigcap_{\lambda_i \in \Lambda_i^*} H_i^p(\lambda_i)$. *Q.E.D.*

Parts (a) and (b) of the above lemma show that the LP problem reduces to computing the maximal score for each regime R in isolation. The next lemma establishes that the set of BFXE payoffs with respect to p is included in the set Q^p.

Lemma 2. *For every $\delta \in (0,1)$, $p \in \triangle\mathscr{R}$, and $i \in I$, $E_i^p(\delta) \subseteq coE_i^p(\delta) \subseteq Q_i^p$. Consequently, $E_i^p(\delta) \subseteq coE_i^p(\delta) \subseteq Q_i^p$.*

The proof is analogous to Theorem 3.1 (i) of Fudenberg and Levine (1994); we provide the formal proof in Appendix for completeness.

4.2 Computing $E(\delta)$ with Patient Players

In the previous subsection, it is shown that the equilibrium payoff set $E^p(\delta)$ is bounded by the set Q^p. Now we prove that this bound is tight when players are patient. As argued by Fudenberg, Levine, and Maskin (1994), when δ is close to one, a small variation of the continuation payoffs is sufficient for incentive provision, so that we can focus on the continuation payoffs w near the target payoff vector v. Based on this observation, we obtain the following lemma, which asserts that "local generation" is sufficient for self-generation with patient players.

Definition 6. A subset W_i of $\mathbb{R}^{|\Omega|}$ is *locally ex-post generating with respect to $p \in \triangle\mathscr{R}$* if for each $v_i \in W_i$, there is a discount factor $\delta_{v_i} \in (0,1)$ and an open neighborhood U_{v_i} of v_i such that $W_i \cap U_{v_i} \subseteq B_i^p(\delta_{v_i}, W_i)$.

Lemma 3. *For each $i \in I$, let W_i be a subset of $\mathbb{R}^{|\Omega|}$ that is compact, convex, and locally ex-post generating with respect to $p \in \triangle\mathscr{R}$. Then there is $\overline{\delta} \in (0,1)$ such that $\times_{i\in I} W_i \subseteq E^p(\delta)$ for all $\delta \in (\overline{\delta}, 1)$.*

Proof. This is a straightforward generalization of Lemma 4.2 of Fudenberg, Levine, and Maskin (1994). *Q.E.D.*

The next lemma shows that the set Q^p is included in the limit set of stationary BFXE payoffs with respect to p.

Definition 7. A subset W_i of $\mathbb{R}^{|\Omega|}$ is *smooth* if it is closed and convex; it has a nonempty interior; and there is a unique unit normal for each point on its boundary.[13]

Lemma 4. *For each $i \in I$, let W_i be a smooth subset of the interior of Q_i^p. Then there is $\overline{\delta} \in (0,1)$ such that for $\delta \in (\overline{\delta}, 1)$, $\times_{i\in I} W_i \subseteq E^p(\delta)$.*

The proof is similar to Theorem 3.1 (ii) of Fudenberg and Levine (1994), and again we give the formal proof in Appendix for completeness. To prove the lemma, we show that a smooth subset W_i is locally ex-post generating; then Lemma 3 applies and we can conclude that W_i is in the equilibrium payoff set when players are patient.

Combining Lemmas 2 and 4, we obtain the next proposition, which asserts that the limit set of stationary BFXE payoffs with respect to p is equal to the set Q^p.

Proposition 5. *If $\dim Q_i^p = |\Omega|$ for each $i \in I$, then $\lim_{\delta\to 1} E^p(\delta) = Q^p$.*

Now we characterize the set of all stationary BFXE payoffs, $E(\delta) = \bigcup_{p\in\triangle\mathscr{R}} E^p(\delta)$, in the limit as $\delta \to 1$. This is a counterpart of Proposition 4 of EHO.

Proposition 6. *Suppose that there is $p \in \triangle\mathscr{R}$ such that $\dim Q_i^p = |\Omega|$ for each $i \in I$. Then $\lim_{\delta\to 1} E(\delta) = \bigcup_{p\in\triangle\mathscr{R}} Q^p$.*

Proof. From Proposition 5, it follows that $\lim_{\delta\to 1} E(\delta) = \bigcup_{p\in\triangle\mathscr{R}} Q^p$ if $\dim Q_i^p = |\Omega|$ for all $i \in I$ and $p \in \triangle\mathscr{R}$. Here we prove that the same conclusion holds if there is $p \in \triangle\mathscr{R}$ such that $\dim Q_i^p = |\Omega|$ for each $i \in I$.

[13] A sufficient condition for each boundary point of W_i to have a unique unit normal is that the boundary of W_i is a C^2-submanifold of $\mathbb{R}^{|\Omega|}$.

Let v_i be an interior point of $\bigcup_{p\in\triangle\mathscr{R}} Q^p$. It suffices to show that there is $p \in \triangle\mathscr{R}$ such that v_i is an interior point of Q^p. Let $\hat{p} \in \triangle\mathscr{R}$ be such that $\dim Q_i^{\hat{p}} = |\Omega|$ for each $i \in I$, and $\hat{v}_i$ be an interior point of $Q^{\hat{p}}$. Since v_i is in the interior of $\bigcup_{p\in\triangle\mathscr{R}} Q^p$, there are $\tilde{v}_i$ and $\kappa \in (0,1)$ such that $\tilde{v}_i$ is in the interior of $\bigcup_{p\in\triangle\mathscr{R}} Q^p$ and $\kappa\hat{v}_i + (1-\kappa)\tilde{v}_i = v_i$. Let $\tilde{p} \in \triangle\mathscr{R}$ be such that $\tilde{v}_i \in Q^{\tilde{p}}$, and let $p \in \triangle\mathscr{R}$ be such that $p = \kappa\hat{p} + (1-\kappa)\tilde{p}$.

We claim that v_i is an interior point of Q^p. From Lemma 1(b),

$$\begin{aligned} k_i^p(\lambda_i) = & \sum_{R\in\mathscr{R}} p(R)k_i^R(\lambda_i) \\ = & \kappa \sum_{R\in\mathscr{R}} \hat{p}(R)k_i^R(\lambda_i) + (1-\kappa) \sum_{R\in\mathscr{R}} \tilde{p}(R)k_i^R(\lambda_i) \\ = & \kappa k_i^{\hat{p}}(\lambda_i) + (1-\kappa)k_i^{\tilde{p}}(\lambda_i) \end{aligned}$$

for all λ_i. Since $\hat{v}_i$ is in the interior of $Q^{\hat{p}}$, we have $k_i^{\hat{p}}(\lambda_i) > \lambda_i \cdot \hat{v}_i$ for all λ_i. Likewise, since $\tilde{v}_i \in Q^{\tilde{p}}$, $k_i^{\tilde{p}}(\lambda_i) \geq \lambda_i \cdot \tilde{v}_i$ for all λ_i. Substituting these inequalities,

$$k_i^p(\lambda_i) > \kappa\lambda_i \cdot \hat{v}_i + (1-\kappa)\lambda_i \cdot \tilde{v}_i = \lambda_i \cdot v_i$$

for all λ_i. This shows that v_i is an interior point of Q^p. *Q.E.D.*

So far we have focused on stationary BFXE and characterized its limit equilibrium payoff set. As mentioned in Section 2.3, the limit set of stationary BFXE payoffs is equal to the limit set of all BFXE payoffs with public randomization; therefore it follows from Proposition 6 that if there is $p \in \triangle\mathscr{R}$ such that $\dim Q_i^p = |\Omega|$ for each $i \in I$, then the limit set of all (possibly non-stationary) BFXE payoffs with public randomization is equal to $\bigcup_{p\in\triangle\mathscr{R}} Q^p$. Moreover, the same result holds even if public randomization is not available, because the set of BFXE payoffs does not depend on the presence of public randomization in the limit as $\delta \to 1$. We omit the formal proof, as it is analogous to the on-line appendix of EHO.

5 State Learning and Belief-Free Equilibria

5.1 General Case

In the last section, we have shown that the limit set of BFXE payoffs is computed by a series of static linear programming problems. In this subsection, we provide a more powerful characterization of the limit equilibrium payoff set, by restricting attention to games that satisfy certain informational conditions. Specifically, we consider games where each player can learn the true state from a sequence of observed signals, and show that the limit set of BFXE payoffs of the overall game equals the product of the limit sets of belief-free equilibrium payoffs of the corresponding known-state games; that is, there are BFXE in which the payoffs are approximately the same as if players learn the true state and play a belief-free equilibrium for that state.

We begin with introducing the informational conditions imposed throughout this subsection. Let $\hat{\pi}^{\omega}_{-i}(a_i,\alpha_{-i}) = (\hat{\pi}^{\omega}_{-i}(a_{-i},\sigma_{-i}|a_i,\alpha_{-i}))_{(a_{-i},\sigma_{-i})}$ denote the probability distribution of (a_{-i},σ_{-i}) when players play (a_i,α_{-i}) at state ω. That is, $\hat{\pi}^{\omega}_{-i}(a_{-i},\sigma_{-i}|a_i,\alpha_{-i}) = \alpha_{-i}(a_{-i})\sum_{\sigma_i\in\Sigma_i}\pi^{\omega}(\sigma_i,\sigma_{-i}|a)$ for each (a_{-i},σ_{-i}). Given an action plan $\vec{\alpha}_{-i}$, let $\Pi^{\omega,R}_{-i}(\vec{\alpha}_{-i})$ be a matrix with rows $\hat{\pi}^{\omega}_{-i}(a_i,\alpha^{R,\theta^{\omega}_{-i}}_{-i})$ for all $a_i \in A_i$. Let $\Pi^{(\omega,\tilde{\omega}),R}_{-i}(\vec{\alpha}_{-i})$ be a matrix constructed by stacking two matrices, $\Pi^{\omega,R}_{-i}(\vec{\alpha}_{-i})$ and $\Pi^{\tilde{\omega},R}_{-i}(\vec{\alpha}_{-i})$.

Definition 8. An action plan $\vec{\alpha}_{-i}$ has *individual full rank for ω at regime R* if $\Pi^{\omega,R}_{-i}(\vec{\alpha}_{-i})$ has rank equal to $|A_i|$. An action plan $\vec{\alpha}_{-i}$ has *individual full rank* if it has individual full rank for all ω and R.

Individual full rank implies that player $-i$ can statistically distinguish player i's deviation using a pair (a_{-i},σ_{-i}) of her action and signal when the true state is ω and the realized public signal is R. Note that this definition is slightly different from those of Fudenberg, Levine, and Maskin (1994) and Fudenberg and Yamamoto (2010b); here we consider the joint distribution of actions and signals, while they consider the distribution of signals.

Definition 9. For each $\omega \in \Omega$, $\tilde{\omega} \neq \omega$, and R, an action plan $\vec{\alpha}_{-i}$ has *statewise full rank for $(\omega,\tilde{\omega})$ at regime R* if $\Pi^{(\omega,\tilde{\omega}),R}_{-i}(\vec{\alpha}_{-i})$ has rank equal to $2|A_i|$.

Statewise full rank assures that player $-i$ can statistically distinguish ω from $\tilde{\omega}$ irrespective of player i's play, given that the realized public signal is R. Again this definition is slightly different from those of Fudenberg and Yamamoto (2010a) and Fudenberg and Yamamoto (2010b), as we consider the joint distribution of actions and signals.

Condition IFR. For each i, every pure action plan $\vec{\alpha}_{-i}$ has individual full rank.

This condition is generically satisfied if there are so many signals that $|\Sigma_{-i}| \geq |A_i|$ for each i. Note that under (IFR), every mixed action plan has individual full rank.

Condition SFR. For each i and $(\omega, \tilde{\omega})$ satisfying $\omega \neq \tilde{\omega}$, either (i) there is $\vec{\alpha}_{-i}$ that has statewise full rank for this pair at some regime $R \in \mathscr{R}$, or (ii) $\theta^{\omega}_{-i} \neq \theta^{\tilde{\omega}}_{-i}$.

This condition (SFR) requires that for each pair $(\omega, \tilde{\omega})$, players can statistically distinguish these two states using (i) observed private signals or (ii) initial private information. Note that (SFR) is sufficient for each player to learn the true state in the long run. For example, suppose that there are three possible states, and that players have no initial information. In periods 1, 4, 7, $\cdots$, let player $-i$ choose an action that has statewise full rank for (ω_1, ω_2) and perform a statistical inference to distinguish ω_1 and ω_2; in periods 2, 5, 8, $\cdots$, let player $-i$ choose an action that has statewise full rank for (ω_2, ω_3) and perform a statistical inference to distinguish ω_2 and ω_3; and in periods 3, 6, 9, $\cdots$, let player $-i$ choose an action that has statewise full rank for (ω_1, ω_3) and perform a statistical inference to distinguish ω_1 and ω_3. When the true state is ω_2, player $-i$ will be likely to accept ω_2 in the first two statistical tests, so that she will believe that the true state is ω_2. Likewise, player $-i$ will eventually learn the true state based on the statistical tests when the true state is ω_1 or ω_3. In this way, under (SFR), each player can learn the true state from a sequence of observed signals (or from initial information). However, note that players do not share any common information, and hence it is unclear if players can coordinate their play state by state.

In what follows, we give a simple characterization of the set of BFXE payoffs when (IFR) and (SFR) hold. Given ω, let G^{ω} denote the infinitely repeated game where players know that the true state is ω, and consider belief-free equilibria of EHO in this known-state game G^{ω}. In particular, given ω and p, let us look at belief-free equilibria where players follow a recommendation by a public signal R with i.i.d. distribution p. (To be more precise, here we consider belief-free equilibria where given a public signal $R = (R_i(\theta_i)_{\theta_i \in \Theta_i})_{i \in I}$, player i chooses her action from the set $R_i(\theta_i^{\omega})$ of actions recommended for state ω, and ignore recommendations for state $\tilde{\omega} \neq \omega$. Note that in these equilibria, signals R and $\tilde{R}$ such that $R_i(\theta_i^{\omega}) = \tilde{R}_i(\theta_i^{\omega})$ yield exactly the same recommended action set.) Let $M_i^{\omega,p}$ and $m_i^{\omega,p}$ be the maximum and minimum of player i's payoffs attained by these belief-free equilibria. As EHO show, these values are calculated by the following formula: $M_i^{\omega,p} = \sup_{\vec{\alpha}_{-i}} M_i^{\omega,p}(\vec{\alpha}_{-i})$ and $m_i^{\omega,p} = \inf_{\vec{\alpha}_{-i}} m_i^{\omega,p}(\vec{\alpha}_{-i})$ where $M_i^{\omega,p}(\vec{\alpha}_{-i})$ is the solution to (LP-Max) and $m_i^{\omega,p}(\vec{\alpha}_{-i})$ is the solution to (LP-Min).

$$\text{(LP-Max)} \qquad \max_{\substack{v_i^\omega \in \mathbb{R} \\ x_i^\omega : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to \mathbb{R}}} v_i^\omega \qquad \text{subject to}$$

$$\text{(i)} \quad v_i^\omega = \sum_{R \in \mathscr{R}} p(R) \sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^\omega(a_i^R, a_{-i}) \\ +\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $(a_i^R)_{R \in \mathscr{R}}$ s.t. $a_i^R \in R_i(\theta_i^\omega)$ for each $R \in \mathscr{R}$,

$$\text{(ii)} \quad v_i^\omega \geq \sum_{R \in \mathscr{R}} p(R) \sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^\omega(a_i^R, a_{-i}) \\ +\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $(a_i^R)_{R \in \mathscr{R}}$ s.t. $a_i^R \in A_i$ for each $R \in \mathscr{R}$,

$$\text{(iii)} \quad x_i^\omega(R, a_{-i}, \sigma_{-i}) \leq 0, \quad \text{for all } R \in \mathscr{R},\ a_{-i} \in A_{-i}, \text{ and } \sigma_{-i} \in \Sigma_{-i}.$$

and

$$\text{(LP-Min)} \qquad \min_{\substack{v_i^\omega \in \mathbb{R} \\ x_i^\omega : A_{-i} \times \Sigma_{-i} \to \mathbb{R}}} v_i^\omega \qquad \text{subject to constraints (i) and (ii) of (LP-Max) and}$$

$$\text{(iii)} \quad x_i^\omega(R, a_{-i}, \sigma_{-i}) \geq 0, \quad \text{for all } R \in \mathscr{R},\ a_{-i} \in A_{-i}, \text{ and } \sigma_{-i} \in \Sigma_{-i}.$$

Note that (LP-Max) is very similar to (LP-Individual) for λ_i such that $\lambda_i^\omega = 1$ and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$. The difference is that constraints (i) and (ii) are imposed not only for ω but also for $\tilde{\omega} \neq \omega$ in (LP-Individual). Also, (LP-Min) is very similar to (LP-Individual) for λ_i such that $\lambda_i^\omega = -1$ and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$.

The main theorem of this subsection is:

Proposition 7. *Suppose that (IFR) and (SFR) hold. Then $Q_i^p = \times_{\omega \in \Omega} [m_i^{\omega,p}, M_i^{\omega,p}]$ for each $i \in I$, $\omega \in \Omega$, and $p \in \triangle \mathscr{R}$. In particular if there is $p \in \triangle \mathscr{R}$ such that $M_i^{\omega,p} > m_i^{\omega,p}$ for all $i \in I$ and $\omega \in \Omega$, then $\lim_{\delta \to 1} E(\delta) = \bigcup_{p \in \triangle \mathscr{R}} \times_{\omega \in \Omega} \times_{i \in I} [m_i^{\omega,p}, M_i^{\omega,p}]$.*

This proposition asserts that if (IFR) and (SFR) hold, then the limit set of BFXE payoffs is isomorphic to the set of maps from states to belief-free equilibrium payoffs. That is, there are BFXE where the payoffs are as if players learn the true state and play a belief-free equilibrium for that state.

To get a deeper understanding of what is going on in the equilibrium strategies, let us focus on BFXE where players are indifferent over all actions in any period and any state.[14] In our equilibria, players' strategy profile has the following properties; (i) player i makes player $-i$ indifferent over all actions given any history and given any state, and (ii) player i controls player $-i$'s payoffs in such a way that player $-i$'s continuation payoffs at state ω is close to the target payoff when player i has learned that the true state is likely to be ω. Property (ii) implies that player i's individual state learning is sufficient for player $-i$'s payoff of the entire game to approximate the target payoffs state by state. Thus, if each player can individually learn the true state, then both players' payoffs approximate the target payoffs state by state (although the state may not necessarily be an approximate common knowledge). Also, this strategy profile satisfies incentive compatibility, since property (i) assures that each player's play is optimal after every history. Note that in the above equilibrium strategies, player i's individual state learning is irrelevant to her own continuation payoffs, and influences player $-i$'s payoffs only. Indeed, it follows from (i) that player i cannot obtain better payoffs by changing her action contingently on what she has learned from the past history.

To prove the proposition, we compute the maximal score of (LP-Individual) for each direction λ_i. We first consider "cross-state" directions λ_i, and prove that under (SFR), the scores for these directions are so high that the maximal half spaces in these directions impose no constraints on the equilibrium payoff set, that is, there is no trade-off between equilibrium payoffs for different states. Specifically, Lemma 5 shows that the maximal scores for cross-state directions are infinitely large if $\vec{\alpha}_{-i}$ has statewise full rank. Lemma 6 shows that if player $-i$ can distinguish states from her initial private information θ_{-i}, then the scores for cross-state directions are so high that the corresponding maximal half spaces do not bound the set Q_i^p. For each $\omega \in \Omega$, let $e(\omega) = (e^{\omega}(\omega))_{\omega\in\Omega} \in \Lambda_i$ be a unit vector such that $e^{\omega}(\omega) = 1$ and $e^{\tilde{\omega}}(\omega) = 0$ for all $\tilde{\omega} \neq \omega$.

Lemma 5. *Suppose that $\vec{\alpha}_{-i}$ has individual full rank, and has statewise full rank for $(\omega, \tilde{\omega})$ at regime R. Then for any p and λ_i satisfying $p(R) > 0$, $\lambda_i^{\omega} \neq 0$, and $\lambda_i^{\tilde{\omega}} \neq 0$. $k_i^p(\vec{\alpha}_{-i}, \lambda_i) = \infty$.*

[14] To be precise, these are stationary BFXE with respect to $p^A \in \triangle\mathscr{R}$, where p^A is the unit vector that puts one to the regime $R = ((R_i(\theta_i))_{\theta_i})_i$ such that $R_i(\theta_i) = A_i$ for all i and θ_i.

This lemma is analogous to Lemma 6 of Fudenberg and Yamamoto (2010a), and we give the formal proof in Appendix for completeness. The main idea is that if $\vec{\alpha}_{-i}$ has statewise full rank for $(\omega, \tilde{\omega})$, then "utility transfer" between ω and $\tilde{\omega}$ can infinitely increase the score.

Lemma 6. *Suppose that (IFR) holds. Let λ_i be such that $\theta_{-i}^{\omega} \neq \theta_{-i}^{\tilde{\omega}}$ for any $(\omega, \tilde{\omega})$ satisfying $\omega \neq \tilde{\omega}$, $\lambda_i^{\omega} \neq 0$, and $\lambda_i^{\tilde{\omega}} \neq 0$. Then $k_i^p(\lambda_i) \geq \sum_{\omega \in \Omega} |\lambda_i^{\omega}| k_i^p(e(\omega))$, and as a result $\bigcap_{\omega \in \Omega} H_i^p(e(\omega)) \subseteq H_i^p(\lambda_i)$.*

The formal proof is given in Appendix, but the intuition is as follows. If player $-i$ can distinguish ω and $\tilde{\omega}$ using private information θ_{-i}, then she can choose different actions contingent on whether the true state is ω or $\tilde{\omega}$. Therefore we expect that the score on state ω will not constrain the score on state $\tilde{\omega}$ so that the maximal score for directions vectors that only weight these two states will be high enough not to constrain the set Q_i^p.

Next we compute the maximal scores for the remaining "single-state" directions. Consider (LP-Individual) for direction λ_i such that $\lambda_i^{\omega} = 1$ for some ω and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$. If (IFR) holds, then there is continuation payoffs that make player i indifferent over all actions, so that constraints (i) and (ii) for $\tilde{\omega} \neq \omega$ are vacuous. Then it turns out that the problem is identical to the one that computes $M_i^{\omega,p}(\vec{\alpha}_{-i})$, and hence we have $k_i^p(\lambda_i) = M_i^{\omega,p}$. Likewise, consider (LP-Individual) for direction λ_i such that $\lambda_i^{\omega} = -1$ for some ω and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$. If (IFR) holds, then the problem is isomorphic to the one that computes $m_i^{\omega,p}(\vec{\alpha}_{-i})$, and as a result we have $k_i^p(\lambda_i) = -m_i^{\omega,p}$. The next lemma summarizes these discussions.

Lemma 7. *Suppose that (IFR) holds. For λ_i such that $\lambda_i^{\omega} = 1$ and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$, $k_i^p(\lambda_i) = M_i^{\omega,p}$. For λ_i such that $\lambda_i^{\omega} = -1$ and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$, $k_i^p(\lambda_i) = -m_i^{\omega,p}$.*

Now we are ready to prove Proposition 7; we use Lemmas 5 through 7 to compute the scores of (LP-Individual) for various directions.

Proof of Proposition 7. From Proposition 6, it suffices to show that $Q_i^p = \times_{\omega \in \Omega}[m_i^{\omega,p}, M_i^{\omega,p}]$ for each i, ω, and p. Let Λ_i^* be the set of all single-state directions, that is, Λ_i^* is the set of all $\lambda_i \in \Lambda_i$ such that $\lambda_i^{\omega} \neq 0$ for some ω and $\lambda_i^{\tilde{\omega}} = 0$ for all $\tilde{\omega} \neq \omega$. Then it follows from Lemmas 5 and Lemma 6 that under (SFR), we have $\bigcap_{\tilde{\lambda}_i \in \Lambda^*} H_i^p(\tilde{\lambda}_i) \subseteq H_i^p(\lambda_i)$ for all $\lambda_i \notin \Lambda_i^*$, so that $Q_i^p = \bigcap_{\lambda_i \in \Lambda} H_i^p(\lambda_i) = \bigcap_{\lambda_i \in \Lambda^*} H_i^p(\lambda_i)$. Note that, from Lemma 7, we have $H_i^p(\lambda_i) = \{v_i \in \mathbb{R}^{|\Omega|} | v_i^\omega \leq M_i^{\omega,p}\}$ for $\lambda_i \in \Lambda_i^*$ such that $\lambda_i^\omega = 1$, and $H_i^p(\lambda_i) = \{v_i \in \mathbb{R}^{|\Omega|} | v_i^\omega \geq m_i^{\omega,p}\}$ for each $\lambda_i \in \Lambda_i^*$ such that $\lambda_i^\omega = -1$. Therefore, $Q_i^p = \bigcap_{\lambda_i \in \Lambda^*} H_i^p(\lambda_i) = \times_{\omega \in \Omega} [m_i^{\omega,p}, M_i^{\omega,p}]$, as desired. *Q.E.D.*

Remark 2. As a corollary of Proposition 7, we can derive a sufficient condition for the existence of BFXE with patient players. That is, there are BFXE if players are patient, (IFR) and (SFR) hold, and there is p such that $M_i^{\omega,p} > m_i^{\omega,p}$ for all i and ω. Note that the last condition "$M_i^{\omega,p} > m_i^{\omega,p}$ for all i and ω" implies that there are belief-free equilibria with respect to p for each state ω.

5.2 Games with (Almost) Observable Actions

In this subsection, we consider games where actions are (almost) observable at each state ω. There are many economic applications that correspond to this model: One example is the price-setting duopoly where a firm's price is public information (so that actions are perfectly observable), and its sales level is private information and follows an unknown distribution. Another example is the case where each player is subject to a small chance of observation errors of the opponent's action and the probabilities of observation errors are unknown.[15]

To consider "almost observable actions," we need to think about how to measure the observability of actions. For this, we generalize the concept of ε-perfectness of the past work (e.g., Ely and Välimäki (2002), Hörner and Olszewski (2006), and Yamamoto (2009)) to our setting. For each $\varepsilon \geq 0$, the signal structure π is *ε-perfect* if for each $i \in I$, $a_i \in A_i$, and $\omega \in \Omega$, there is a partition of Σ_i into $\{\Sigma_i(a,\omega)\}_{a_{-i} \in A_{-i}}$ such that $\sum_{\Sigma_i(a,\omega)} \pi_i^\omega(\sigma_i|a) \geq 1-\varepsilon$ for each $a_{-i} \in A_{-i}$. In words, given any action profile $a \in A$ and state ω, the probability that player i observes a signal from the corresponding set $\Sigma_i(a,\omega)$ is at least $1-\varepsilon$. Note that ε-perfectness does not impose any restriction on the observability of the true state ω.

[15]Wiseman (2010) provides a sufficient condition for the folk theorem for games where actions are perfectly observable and players observe public and private signals about the state of the world in every period. The key condition for his folk theorem is that the distribution of public signals depends on the state so that players can learn the true state from public signals. In our setting, players need to use private signals to learn the true state, so that his folk theorem does not apply.

By definition, every action profile has individual full rank if the signal structure is ε-perfect with ε sufficiently close to 0. Also we assume that the distribution of private signals depends on the state so that (SFR) holds. Then Proposition 7 applies, and the limit equilibrium payoff set is $\times_{i\in I}\times_{\omega\in\Omega}[m_i^{\omega,p}, M_i^{\omega,p}]$.

Let $\pi = (\pi^{\omega}(\cdot|a))_{(\omega,a)}$ denote the signal structure, and we write $M_i^{\omega,p}(\pi)$ and $m_i^{\omega,p}(\pi)$ when we want to emphasize the dependence on the signal structure. As Lemma 3 of EHO shows, $M_i^{\omega,p}(\pi)$ and $m_i^{\omega,p}(\pi)$ are monotonic with respect to π (equipped with the Blackwell ordering), and they have the following limits when the signal structure converges to the 0-perfect signal structure:

$$\overline{M}_i^{\omega,p} = \sum_{R\in\mathscr{R}} p(R) \max_{\alpha_{-i}\in\triangle R_{-i}(\theta_{-i}^{\omega})} \min_{a_i\in R_i(\theta_i^{\omega})} g_i^{\omega}(a_i,\alpha_{-i}) \tag{3}$$

and

$$\overline{m}_i^{\omega,p} = \sum_{R\in\mathscr{R}} p(R) \min_{\alpha_{-i}\in\triangle R_{-i}(\theta_{-i}^{\omega})} \max_{a_i\in A_i} g_i^{\omega}(a_i,\alpha_{-i}). \tag{4}$$

In words, $\overline{M}_i^{\omega,p}$ is player i's worst payoff at state ω, given that player $-i$ tries to reward player i, and given that players have to choose actions from a recommended set. Likewise, $\overline{m}_i^{\omega,p}$ is player i's maximum payoff at ω, given that player $-i$ tries to punish player i, and given that player $-i$ has to choose an action from a recommended set.

From the above observations, we obtain the following proposition:

Proposition 8. *For any $\eta > 0$, there is $\varepsilon > 0$ such that $\times_{\omega\in\Omega}[\overline{m}_i^{\omega,p}+\eta, \overline{M}_i^{\omega,p}-\eta] \subseteq Q_i^p$ for any $p\in\triangle\mathscr{R}$, for any $i\in I$, and for any ε-perfect signal structure π satisfying (SFR). In particular, given any 0-perfect signal structure π satisfying (SFR), $\times_{\omega\in\Omega}[\overline{m}_i^{\omega,p}, \overline{M}_i^{\omega,p}] = Q_i^p$.*

5.3 Applications

Now we apply Proposition 8 to several examples, and show that there are asymptotically efficient equilibria. This highlights that our state-learning theorem is useful especially when actions are (almost) perfectly observable and players' payoffs are private information and follow an unknown distribution.

Example 1. *Investment Games with Positive Externalities and Free Riding.* There are two players and two possible states, i.e., $\Omega = \{\omega_1, \omega_2\}$. In each period, each player makes a decision on whether investing or not. So let $A_i = \{C_i, D_i\}$ for each i, where C_i implies investing and D_i implies not investing. After making decisions, each player i receives a stochastic output z_i from a finite set Z_i, where z_i is private information and its distribution depends on the true state ω and on the total investment $a \in A$. We assume that investments are perfectly observable; thus the set of player i's signals is $\Sigma_i = A \times Z_i$, and $\pi^\omega(\sigma|a) = 0$ for each ω, a, and $\sigma = (\sigma_1, \sigma_2) = ((a', z_1), (a'', z_2))$ such that $a' \neq a$ or $a'' \neq a$. Player i's actual payoff does not depend on the state ω and has a form $u_i(a_i, \sigma_i) = \tilde{u}_i(z_i) - c_i(a_i)$, where $\tilde{u}_i(z_i)$ is player i's profit from an output z_i and $c_i(a_i)$ is cost of investments; let $c_i(a_i) = \tilde{c}_i > 0$ for $a_i = C_i$ and $c_i(a_i) = 0$ for $a_i = D_i$. The expected payoff of firm i at state ω is denoted by $g_i^\omega(a) = \sum_{\sigma \in \Sigma} \pi^\omega(\sigma|a) u_i(a_i, z_i)$. Let $\Theta_i = \{\Omega\}$ for each i, that is, players do not the true distribution of z_i and hence do not know the marginal profit from investing.

We assume that a player's investment has a *positive externality* to the distribution of the opponent's output, and hence increases the opponent's expected payoff. That is, $g_i^\omega(a_i, C_{-i}) > g_i^\omega(a_i, D_{-i})$ for each i, ω, and a_i. Also, investments are socially efficient, so that $g_1^\omega(C_i, C_{-i}) + g_2^\omega(C_i, C_{-i}) \geq g_1^\omega(a) + g_2^\omega(a)$ for each ω and a. However, a player has *free-riding incentives* in that $g_i^\omega(D_i, C_{-i}) > g_i^\omega(C_i, C_{-i})$ for each i and ω, and hence the efficient outcome (C_1, C_2) is not a static equilibrium at both states.

The following economic examples fit this model.

- Intra-brand competition. There are two independent retailers sharing the same brand name (distributing one manufacturer's output etc.), and they decide whether promoting the brand name (C_i) or not promoting (D_i). Sales level z_i of retailer i is stochastic and its distribution is unknown. The promotion activity by a retailer increases its own expected sales level and the opponent's expected sales level; thus each distributor can free-ride, and benefit from the promotional activity of the other distributor, as pointed out by Telser (1960).

- Lobbying. There are two firms in the same industry, and they decide whether lobbying (C_i) or not lobbying (D_i). The excise tax rate is determined by the total amount of lobbying contributions, and then firms sell the goods. Firm i's profit z_i from selling the goods is stochastic and its distribution is unknown. As in Pecorino (1998), each firm can benefit from lobbing by the rival firm, as it lowers the excise tax rate and thereby gives more expected profits to each firm.

- Common space or environment. Two stores are located on the same street, and each store chooses whether doing street maintenance (C_i) or not (D_i). Sales level (or the number of customers) z_i at store i is stochastic, and likely to be high if the street is clean (but the exact distribution is unknown). Each store benefits if the other store keeps the street clean, and may have free-riding incentives.

We further assume that when the opponent does not invest, the marginal profit from investing is less than the cost at state ω_1, but is more than the cost at state ω_2. That is, $g_i^{\omega_1}(D_i,D_{-i}) > g_i^{\omega_1}(C_i,D_{-i})$ and $g_i^{\omega_2}(C_i,D_{-i}) > g_i^{\omega_2}(D_i,D_{-i})$ for each i.[16] Under this assumption, the stage game at state ω_1 is the prisoner's dilemma (i.e., $g_i^{\omega_1}(D_i,C_{-i}) > g_i^{\omega_1}(C_i,C_{-i}) > g_i^{\omega_1}(D_i,D_{-i}) > g_i^{\omega_1}(C_i,D_{-i})$), while the stage game at state ω_2 is a chicken game (i.e., $g_i^{\omega_2}(D_i,C_{-i}) > g_i^{\omega_2}(C_i,C_{-i}) > g_i^{\omega_2}(C_i,D_{-i}) > g_i^{\omega_2}(D_i,D_{-i})$). Therefore, there is no static ex-post equilibrium in this example. Note that the folk theorems of Fudenberg and Yamamoto (2010a) and Wiseman (2010) do not apply here, as they assume that players obtain public (or almost public) information about the true state in each period.

Now we show that there are BFXE where the payoffs approximate efficiency state by state. Note that (SFR) is satisfied for generic signal distributions. Consider $p \in \triangle\mathscr{R}$ such that $p(A) = 1$ and $p(R) = 0$ for other R. Using (3) and (4), it follows that $\overline{M}_i^{p,\omega} = g_i^{\omega}(C_i,C_{-i})$ and $\overline{m}_i^{p,\omega} = \max_{a_i \in A_i} g_i^{\omega}(a_i,D_{-i}) < g_i^{\omega}(C_i,C_{-i})$ for each i and ω. This implies that Q^p is full dimensional and contains the payoff vector of the profile (C_1,C_2); then Proposition 8 implies that the payoff vector of the profile (C_1,C_2) is in the limit set of BFXE payoffs. That is, efficiency is achieved even if players do not know the marginal profit from investing.[17]

[16] If $g_i^{\omega}(D_i,D_{-i}) \geq g_i^{\omega}(C_i,D_{-i})$ for each i and ω, then the stage game is the prisoner's dilemma at each state, and hence (D_1,D_2) is a static ex-post equilibrium. In this case the efficient outcome is trivially achieved by a simple trigger strategy. The same thing happens if $g_i^{\omega}(D_i,D_{-i}) \leq g_i^{\omega}(C_i,D_{-i})$ for each i and ω.

[17] A similar result applies to the following price-setting duopoly market. Let $\Omega = \{\omega_1,\omega_2\}$ and $A_i = \{\overline{a}_i, \underline{a}_i\}$, where, $\overline{a}_i$ is the high price, and $\underline{a}_i$ is the low price. Let Z_i be a finite set of possible sales levels of player i, and let $\Sigma_i = A \times Z_i$ be the set of player i's private signals, We assume that $\pi^{\omega}(\sigma|a) = 0$ for each ω, a, and $\sigma = (\sigma_1,\sigma_2) = ((a',z_1),(a'',z_2))$ such that $a' \neq a$ or $a'' \neq a$. That

Example 2. *Conflicting Interests at Different States.* Suppose that there are two states and that players have no initial private information, i.e., $\Omega = \{\omega_1, \omega_2\}$ and $\Theta_i = \{\Omega\}$. Player 1 chooses an action from $A_1 = \{U, D\}$ while player 2 chooses an action from $A_2 = \{L, R\}$. We assume that actions are perfectly observable but the state is not. Specifically, player i's signal set is $\Sigma_i = A \times Z_i$ and $\pi_i^\omega(\sigma_i|a) = 0$ for any $\sigma_i = (\tilde{a}, z_i)$ such that $a \neq \tilde{a}$. Assume that given any pure action profile a, the marginal distribution of z_i at state ω_1 is different from the one at state ω_2, so that (SFR) holds. The payoffs for state ω_1 are shown in the left panel, and those for state ω_2 in the right.

	L	R
U	1, 1	−1, 2
D	2, −1	0, 0

	L	R
U	0, 0	2, −1
D	−1, 2	1, 1

Note that the stage game is the prisoner's dilemma for each state, but the role of actions are reversed; (U, L) is efficient at state ω_1 while (D, R) is efficient at state ω_2. Since the efficient action profiles are different at different states, the efficient payoff vector $((1,1),(1,1))$ is not attainable in one-shot stage game. Nevertheless

is, price is perfectly observable (so that this is not a secret price-cutting game), and the state may affect the distribution of sales levels (z_1, z_2). Firm i's actual payoff does not depend on the state ω and has a form $u_i(a_i, \sigma_i) = a_i z_i - c_i(z_i)$, where $c_i(z_i)$ is cost of production. Assume that $\Theta_i = \{\Omega\}$ for each i, so that firms do not know the true state.

Suppose that there are 100 potential customers in the market, and that these customers have heterogeneous preferences as in Varian (1980). Specifically, we consider the following setting:

- There are d_0 customers who purchase a good with probability $\beta > 0$ at the firm with the cheapest price in each period. They purchase a good at each firm equally likely (i.e., with probability $\frac{\beta}{2}$ and $\frac{\beta}{2}$) if both firms choose the same price.
- For each $i \in I$, there are d_i customers who purchase a good with probability β at firm i (irrespective of its price) in each period.
- There are d_3 customers who never purchase a good if both firms choose the high price. If one firm chooses the low price, then they purchase a good at that firm with probability β. If both firms choose the low price, then they purchase a good at each firm equally likely.

Suppose that $(d_0, d_1, d_2, d_3) = (30, 10, 10, 50)$ at state ω_1, and $(d_0, d_1, d_2, d_3) = (10, 20, 20, 50)$ at state ω_2. Also, let $\overline{a}_i = 4$, $\underline{a}_i = 2$, and $c_i(z_i) = z_i$ for each i. Then, the stage game at state ω_1 is the prisoner's dilemma (note that $g_i^{\omega_1}(\underline{a}_i, \overline{a}_{-i}) = 90\beta$, $g_i^{\omega_1}(\overline{a}_i, \overline{a}_{-i}) = 75\beta$, $g_i^{\omega_1}(\underline{a}_i, \underline{a}_{-i}) = 50\beta$, and $g_i^{\omega_1}(\overline{a}_i, \underline{a}_{-i}) = 30\beta$), while the stage game at state ω_2 is a chicken game (note that $g_i^{\omega_2}(\underline{a}_i, \overline{a}_{-i}) = 80\beta$, $g_i^{\omega_2}(\overline{a}_i, \overline{a}_{-i}) = 75\beta$, $g_i^{\omega_2}(\overline{a}_i, \underline{a}_{-i}) = 60\beta$, and $g_i^{\omega_2}(\underline{a}_i, \underline{a}_{-i}) = 50\beta$). Note that the mixed action α_{-i} with $\alpha_{-i}(\overline{a}_{-i}) = \alpha_{-i}(\underline{a}_{-i}) = \frac{1}{2}$ has a statewise full rank, so that (SFR) follows. Thus there are BFXE where the payoffs approximate the payoff vector of $(\overline{a}_1, \overline{a}_2)$. That is, the cartel is self-enforcing even if the firms do not know the distribution of sales levels.

we will show that players can approximate efficiency in the long run, as each player can learn the true state and condition their play on the state.

Consider $p \in \triangle\mathscr{R}$ such that $p(A) = 1$ and $p(R) = 0$ for other R. It follows from (3) and (4) that $\overline{M}_i^{\omega,p} = 1$ and $\overline{m}_i^{\omega,p} = 0$ for all i and ω. Therefore, $Q_i^p = \times_{\omega\in\Omega}[0,1]$, so that the efficient payoff $((1,1),(1,1))$ is in the set Q^p. Since Q_i^p is full dimensional, Proposition 8 applies and we conclude that $((1,1),(1,1))$ is in the limit equilibrium payoff set. This example shows that BFXE can often approximate efficiency even if the efficient action profiles are different at different states,

Example 3. *Prisoner's Dilemma with Observation Errors.* Here we consider the prisoner's dilemma with a small chance of observation errors, where the probability of observation errors is unknown. When the signal distribution is known, Sekiguchi (1997), Bhaskar and Obara (2002), Piccione (2002), and Ely and Välimäki (2002) show that the efficient outcome is approximated in the prisoner's dilemma with almost-perfectly observable actions. We extend this efficiency result to the case of an unknown signal distribution.

Suppose that there are two states and players have no initial private information. Each player i chooses an action from $A_i = \{C_i, D_i\}$, and player i's signal set is Σ_i with $|\Sigma_i| \geq 2$ for each i. We assume that the signal structure depends on the true state ω, and is ε-perfect. Note that (SFR) generically holds (for example, a mixed action α_{-i} with $\alpha_{-i}(C_{-i}) = \alpha_{-i}(D_{-i}) = \frac{1}{2}$ has statewise full rank generically). Suppose that payoffs are such that the stage game is the prisoner's dilemma for both states; i.e., (C_1, C_2) is efficient while (D_1, D_2) is a static equilibrium at each state.

Consider $p \in \triangle\mathscr{R}$ such that $p(A) = 1$ and $p(R) = 0$ for other R. Note that $\overline{M}_i^{\omega,p} = g_i^\omega(C_i, C_{-i})$ and $\overline{m}_i^{\omega,p} = g_i^\omega(D_i, D_{-i})$ for all $i \in I$ and $\omega \in \Omega$. Therefore, Q_i^p approximates the set $\times_{\omega\in\Omega}[g_i^\omega(D_i, D_{-i}), g_i^\omega(C_i, C_{-i})]$ in the limit as $\varepsilon \to 0$. In particular, the payoff vector of (C_1, C_2) is in the limit set Q^p; that is, the efficient payoff vector can be approximated in the limit as $\varepsilon \to 0$.

6 BFXE and Review Strategies

In repeated games with private monitoring and with a known state, several papers combine the idea of review strategies and belief-free equilibria to attain a larger payoff set than that of belief-free equilibria (Matsushima (2004), EHO, Yamamoto (2007), and Yamamoto (2010)). This method works well especially for games with *independent monitoring*, where players observe statistically independent signals conditional on an action profile and an unobservable common shock: For example, the folk theorem is established for the repeated prisoner's dilemma with independent monitoring.

The idea of review strategies is roughly as follows. The infinite horizon is regarded as a sequence of review phases with length T. Within a review phase, players play the same action and pool the private signals. After a T-period play, the pooled private signals are used to test whether the opponent deviated or not. The law of large numbers assures that this statistical test has an arbitrarily high power; that is, a player can obtain precise information about the opponent's action. The past work constructs a review-strategy equilibrium such that a player's play is belief-free at the beginning of each review phase, assuming that the signal distribution is conditionally independent. Under conditionally independent monitoring, a player's private signals within a review phase does not have any information about whether she could "pass" the opponent's statistical test, which greatly simplifies the verification of the incentive compatibility.

In this section, we show that this idea can be extended to the case where players do not know the true state, although the constructive proof of the past work does not directly apply. Specifically, we consider review strategies where a player's play is belief-free and ex-post optimal at the beginning of each T-period review phase, and compute its equilibrium payoff set. In Section 6.1, we extend the static LP problem of Section 4 to T-period LP problems, and establish that the intersection of the corresponding hyperplanes is the limit set of review-strategy equilibrium payoffs. Then in Section 6.2, we obtain a sharp characterization of the limit equilibrium payoff set, by restricting attention to games where the signal distribution is conditionally independent and players can learn the true state from private signals in the long run. We apply this result is to a secret price-cutting game, and show that cartel is self-enforcing even if firms do not know how profitable the market is. So, readers who are interested in the state-learning theorem and its application might want to skip Section 6.1 and go to Section 6.2 directly.

6.1 Linear Programming Problems for Review Strategies

In this subsection, we consider T-period LP problems as an extension of (LP-Individual), and show that the intersection of the corresponding hyperplanes is the set of review-strategy equilibrium payoffs. Kandori and Matsushima (1998) also consider T-period LP problems to characterize the equilibrium payoff set for repeated games with private monitoring and communication, but our result is not a straightforward generalization of theirs and requires a new proof technique. We elaborate this point in Remark 3 below.

Let S_i^T be the set of player i's strategies for a T-period repeated game, that is, S_i^T is the set of all $s_i^T : \bigcup_{t=0}^{T-1} H_i^t \to \triangle A_i$. Let $\pi_{-i}^{T,\omega}(a)$ denote the distribution of private signals $(\sigma_{-i}^1, \cdots, \sigma_{-i}^T)$ in a T-period repeated game at state ω when players choose the action profile a for T periods; that is, $\pi_{-i}^{T,\omega}(\sigma_{-i}^1, \cdots, \sigma_{-i}^T|a) = \prod_{t=1}^T \pi_{-i}^\omega(\sigma_{-i}^t|a)$. Also, let $\pi_{-i}^{T,\omega}(s_i^T, a_{-i})$ denote the distribution of $(\sigma_{-i}^1, \cdots, \sigma_{-i}^T)$ when player $-i$ chooses action a_{-i} for T periods but player i plays $s_i^T \in S_i^T$. Let $g_i^{T,\omega}(s_i^T, a_{-i}, \delta)$ denote player i's average payoff for a T-period repeated game at state ω, when player i plays s_i^T and player $-i$ chooses a_{-i} for T periods. Throughout this section, we maintain the following assumption.

Definition 10. The signal distribution has *full support* if $\pi^\omega(\sigma|a) > 0$ for all $\omega \in \Omega$, $a \in A$, and $\sigma \in \Sigma$.

As Sekiguchi (1997) shows, if the signal distribution has full support, then for any Nash equilibrium $s \in S$, there is a sequential equilibrium $\tilde{s} \in S$ that yields the same outcome. Therefore, the set of sequential equilibrium payoffs is identical with the set of Nash equilibrium payoffs.

In Section 4, we consider LP problems where one-shot game is played and player i receives a sidepayment x_i^ω contingent on the opponent's history of the one-shot game. In this section we consider LP problems where a T-period repeated game is played and player i receives a sidepayment x_i^ω contingent on the opponent's T-period history. In particular, we are interested in a situation where players perform an action plan profile $\vec{\alpha}$ in the first period (i.e., players observe a public signal $R \in \mathscr{R}$ with distribution $p \in \triangle\mathscr{R}$ before play begins, and choose a possibly mixed action from a recommended set in the first period) and then in the second or later period, players play the pure action chosen in the first period. Also we assume that x_i^ω depends on h_{-i}^T only though the initial public signal, player $-i$'s action in period one, and the sequence of player $-i$'s private signals from period one to period T; that is, a sidepayment to player i at state ω is denoted by $x_i^\omega(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)$. In this scenario, player i's expected overall payoff at state ω (i.e., the sum of the average stage-game payoffs of the T-period repeated game and the sidepayment) when player i chooses an action a_i is equal to

$$\begin{aligned}
&\sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha^{R,\theta_{-i}^\omega}(a_{-i}) \left[\frac{1-\delta}{1-\delta^T} \sum_{t=1}^{T} \delta^{t-1} g_i^\omega(a) + \pi_{-i}^{T,\omega}(a) \cdot x_i^\omega(R, a_{-i}) \right] \\
&= \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha^{R,\theta_{-i}^\omega}(a_{-i}) \left[g_i^\omega(a) + \pi_{-i}^{T,\omega}(a) \cdot x_i^\omega(R, a_{-i}) \right],
\end{aligned}$$

where $x_i^\omega(R, a_{-i}) = (x_i^\omega(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T))_{(\sigma_{-i}^1, \cdots, \sigma_{-i}^T)}$. Here, note that $\pi_{-i}^{T,\omega}(a)$ denotes the distribution of $(\sigma_{-i}^1, \cdots, \sigma_{-i}^T)$ at state ω when the profile a is played for T periods, and the term $\pi_{-i}^{T,\omega}(a) \cdot x_i^\omega(R, a_{-i})$ is the expected sidepayment when the initial public signal is R and the profile a is played for T periods.

Now we introduce the T-period LP problem. For each $(T, \vec{\alpha}_{-i}, \lambda_i, \delta, K)$ where $K > 0$, let $k_i^p(T, \vec{\alpha}_{-i}, \lambda_i, \delta, K)$ be a solution to the following problem:

(T-LP) $\displaystyle\max_{\substack{v_i \in \mathbb{R}^{|\Omega|} \\ x_i : \mathscr{R} \times A_{-i} \times (\Sigma_{-i})^T \to \mathbb{R}^{|\Omega|}}} \lambda_i \cdot v_i$ subject to

(i) $$v_i^\omega = \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^\omega(a_i^R, a_{-i}) \\ +\pi_{-i}^{T,\omega}(a_i^R, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $\omega \in \Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ s.t. $a_i^R \in R_i(\theta_i^\omega)$ for each $R \in \mathscr{R}$,

(ii) $$v_i^\omega \geq \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} g_i^{T,\omega}(s_i^{T,R}, a_{-i}, \delta) \\ +\pi_{-i}^{T,\omega}(s_i^{T,R}, a_{-i}) \cdot x_i^\omega(R, a_{-i}) \end{array} \right]$$

for all $\omega \in \Omega$ and $(s_i^{T,R})_{R\in\mathscr{R}}$ s.t. $s_i^{T,R} \in S_i^T$ for each $R \in \mathscr{R}$,

(iii) $\lambda_i \cdot x_i(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) \leq 0$

for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$, and $(\sigma_{-i}^1, \cdots, \sigma_{-i}^T) \in (\Sigma_{-i})^T$.

(iv) $|x_i(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)| \leq K$

for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$, and $(\sigma_{-i}^1, \cdots, \sigma_{-i}^T) \in (\Sigma_{-i})^T$.

Constraint (i) implies adding-up, that is, the target payoff v_i is exactly achieved if player i chooses an action from the recommended set in the first period and plays the same action until period T. Constraint (ii) is incentive compatibility, that is, player i is willing to choose her action from the recommended set and to play the same action until period T. Constraint (iii) says that a payment x_i lies in the half-space corresponding to direction λ_i. Note that constraints (i) through (iii) of (T-LP) are similar to those of (LP-Individual). Constraint (iv) has not appeared in (LP-Individual), and is new to the literature, as explained in Remark 3 below. This new constraint requires a payment x_i to be bounded by some parameter K.

Recall that the score $k_i^p(\vec{\alpha}_{-i}, \lambda_i, \delta)$ of (LP-Individual) does not depend on δ, as δ does not appear in (LP-Individual). It maybe noteworthy that the same technique does not apply to (T-LP). To see this, note that player i's average payoff $g_i^{T,\omega}(s_i^{T,R}, a_{-i}, \delta)$ of the T-period interval depends on δ when player i plays a non-constant action. Then a pair (v_i, x_i) that satisfies constraint (ii) for some δ may not satisfy constraint (ii) for $\tilde{\delta} \neq \delta$. Therefore the score of (T-LP) may depend on δ.[18]

Let

$$\begin{aligned}
k_i^p(T, \lambda_i, \delta, K) &= \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} k_i^p(T, \vec{\alpha}_{-i}, \lambda, \delta, K), \\
k_i^p(T, \lambda_i, K) &= \liminf_{\delta \to 1} k_i^p(T, \lambda_i, \delta, K), \\
k_i^p(T, \lambda_i) &= \lim_{K \to \infty} k_i^p(T, \lambda_i, K), \\
H_i^p(T, \lambda_i) &= H_i(\lambda_i, k_i^p(T, \lambda_i)),
\end{aligned}$$

and

$$Q_i^p(T) = \bigcap_{\lambda_i \in \Lambda_i} H_i^p(T, \lambda_i).$$

Here $k_i^p(T, \lambda_i, K)$ is defined to be the limit inferior of $k_i^p(T, \lambda_i, \delta, K)$, since $k_i^p(T, \lambda_i, \delta, K)$ may not have a limit as $\delta \to 1$. On th other hand $k_i^p(T, \lambda_i, K)$ has a limit as $K \to \infty$, since $k_i^p(T, \lambda_i, K)$ is increasing with respect to K.

[18] Note that the new constraint (iv) is not an issue here; indeed, it is easy to check that even if we add (iv) to the set of constraints of (LP-Individual) the score of the new LP problem does not depend on δ.

The next proposition is a counterpart to Lemma 4; it shows that the set $\times_{i\in I}Q_i^p(T)$ is a subset of the set of sequential equilibrium payoffs. Note that here we do not assume the signal distribution to be conditionally independent.

Proposition 9. *Suppose that the signal distribution has full support. Let T and p be such that* $\dim Q_i^p(T) = |\Omega|$ *for each $i \in I$. Then the set $\times_{i\in I}Q_i^p(T)$ is in the limit set of sequential equilibrium payoffs as $\delta \to 1$.*

The proof is given in Appendix. In the proof, we (implicitly) show that for any payoff $v \in \times_{i\in I}Q_i^p(T)$, there is a sequential equilibrium with payoff v and such that a player's play is belief-free and ex-post optimal at the beginning of each review phase with length T (while actions in other periods are not necessarily belief-free or ex-post optimal). That is, here we consider *periodically belief-free* and *periodically ex-post* equilibria.[19] Note that the proof of this proposition is not a straightforward generalization of Lemma 4, because δ appears in constraint (ii) of (T-LP). See the following remark for more discussions.

Remark 3. Kandori and Matsushima (1998) also consider T-period LP problems to characterize the equilibrium payoff set for games with private monitoring and communication, but our result is not a mere adaptation of theirs. A main difference is that Kandori and Matsushima (1998) impose "uniform incentive compatibility," which requires the payment scheme to satisfy incentive compatibility for all $\tilde{\delta} \in [\delta, 1)$. They show that with this strong version of incentive compatibility, the local decomposability condition is sufficient for a set W to be self-generating for high δ as in Fudenberg and Levine (1994). On the other hand, our LP problem does not impose uniform incentive compatibility, so that a payment scheme x that satisfies the incentive compatibility constraint (ii) for δ may not satisfy (ii) for $\tilde{\delta} \in (\delta, 1)$. Due to this failure of monotonicity, the local decomposability condition is not sufficient for a set W to be self-generating. Instead, we use the fact that the uniform decomposability condition of Fudenberg and Yamamoto (2010c) is sufficient for a set W to be self-generating. The uniform decomposability condition requires the continuation payoffs w to be within $(1-\delta)K$ of the target

[19] Precisely speaking, in these equilibria, a player's play at the beginning of each review phase is *strongly belief-free* in the sense of Yamamoto (2010); that is, a player's play is optimal regardless of the opponent's past history *and regardless of the opponent's current action*. Indeed, constraints (i) and (ii) of (T-LP) imply that player i's play is optimal given any realization of a_{-i}.

payoff $v \in W$ for all δ, and to prove this property we use the new constraint (iv). Our new LP problem is tractable in the following analysis, as we need to check the incentive compatibility only for a given δ. Note also that the side payment scheme x constructed in the proof of Lemma 8 satisfies constraints (i) through (iv) of (T-LP) but does not satisfy the uniform incentive compatibility of Kandori and Matsushima (1998).

Remark 4. In (T-LP) we restrict attention to the situation where players play the same action throughout the T-period interval, but this is not necessary. That is, even if we consider a LP problem where players play a more complex T-period strategy, we can obtain a result similar to Proposition 9.

6.2 State Learning and Review Strategies

In this subsection, we establish that if the signal distribution satisfies some additional conditions, then there are sequential equilibria where the payoffs are as if players learn the true state and play a *belief-free review-strategy equilibrium* for that state. Belief-free review-strategy equilibria were first considered by Matsushima (2004), who established the folk theorem for the prisoner's dilemma, and extended to general games by EHO, Yamamoto (2007), and Yamamoto (2010). These papers assume the signal distribution to be *weakly conditionally independent*, i.e., players' signal distributions are statistically independent conditional on an action profile a and an unobservable common shock σ_0.[20] In this subsection, we impose the same assumption on the signal distribution:

Condition Weak-CI. There is a finite set Σ_0, $\tilde{\pi}_0^\omega : A \to \triangle\Sigma_0$ for each ω, and $\tilde{\pi}_i^\omega : A \times \Sigma^0 \to \triangle\Sigma_i$ for each (i,ω) such that the following properties hold.

(i) For each $\omega \in \Omega$, $a \in A$, and $\sigma \in \Sigma$,

$$\pi^\omega(\sigma|a) = \sum_{\sigma_0 \in \Sigma_0} \tilde{\pi}_0^\omega(\sigma_0|a) \prod_{i \in I} \tilde{\pi}_i^\omega(\sigma_i|a,\sigma_0).$$

(ii) For each $i \in I$, $\omega \in \Omega$, and $a_{-i} \in A_{-i}$, $\text{rank}\tilde{\Pi}_{-i}^\omega(a_{-i}) = |A_i| \times |\Sigma_0|$ where $\tilde{\Pi}_{-i}^\omega(a_{-i})$ is a matrix with rows $(\tilde{\pi}_{-i}^\omega(\sigma_{-i}|a_i,a_{-i},\sigma_0))_{\sigma_{-i}\in\Sigma_{-i}}$ for all $a_i \in A_i$ and $\sigma_0 \in \Sigma_0$.

[20] Sugaya (2010a) construct belief-free review-strategy equilibria without conditional independence, but he assumes that there are at least four players.

Clause (i) says that the signal distribution is weakly conditionally independent, that is, after players choose profile a, an unobservable common shock σ_0 is randomly selected, and then players observe statistically independent signals conditional on (a,σ_0). Here $\tilde{\pi}_0^\omega(\cdot|a)$ is the distribution of a common shock σ_0 conditional on a, while $\tilde{\pi}_i^\omega(\cdot|a,\sigma_0)$ is the distribution of player i's private signal σ_i conditional on (a,σ_0). Clause (ii) is a strong version of individual full rank; i.e., it implies that player $-i$ can statistically distinguish player i's action a_i and a common shock σ_0. Note that clause (ii) is satisfied generically if $|\Sigma_{-i}| \geq |A_i| \times |\Sigma_0|$ for each i. Note also that (Weak-CI) implies (IFR).

Let $N_i^{\omega,p}$ be the maximum of belief-free review-strategy equilibrium payoffs for the known-state game corresponding to the state ω. Likewise, let $n_i^{\omega,p}$ be the minimum of belief-free review-strategy equilibrium payoffs. As EHO and Yamamoto (2010) show, if the signal distribution is weakly conditionally independent, then these values are calculated by the following formulas:

$$N_i^{\omega,p} = \sum_{R\in\mathscr{R}} p(R) \max_{a_{-i}\in R_{-i}(\theta_{-i}^\omega)} \min_{a_i\in R_i(\theta_i^\omega)} g_i^\omega(a),$$
$$n_i^{\omega,p} = \sum_{R\in\mathscr{R}} p(R) \min_{a_{-i}\in R_{-i}(\theta_{-i}^\omega)} \max_{a_i\in A_i} g_i^\omega(a).$$

Note that these formulas are similar to those of $\overline{M}_i^{\omega,p}$ and $\overline{m}_i^{\omega,p}$ in Section 5.2, but here we do not allow player $-i$ to randomize actions.

The next proposition is the main result in this section; it shows that if the signal distribution is weakly conditionally independent and if each player can learn the true state from observed signal distributions, then there are sequential equilibria where the payoffs are as if players learn the true state and play a belief-free review-strategy equilibrium for that state.

Proposition 10. *Suppose that the signal distribution has full support, and that (SFR) and (Weak-CI) hold. Suppose also that there is $p \in \triangle\mathscr{R}$ such that $N_i^{\omega,p} > n_i^{\omega,p}$ for all i and ω. Then $\bigcup_{p\in\triangle\mathscr{R}} \times_{i\in I} \times_{\omega\in\Omega} [n_i^{\omega,p}, N_i^{\omega,p}]$ is in the limit set of sequential equilibrium payoffs as $\delta \to 1$.*

To prove the proposition, we compute the scores of (T-LP) for various directions. The next two lemmas are extensions of Lemmas 5 and 6, which assert that under (IFR) and (SFR), the scores of (T-LP) for cross-state directions are so high that the half-spaces for these directions impose no restriction on the set $Q_i^p(T)$.

Note that these lemmas do not require the signal distribution to be weakly conditionally independent. The proof of Lemma 8 is found in Appendix.

Lemma 8. *Suppose that (IFR) holds. Suppose also that $\vec{\alpha}_{-i}$ has individual full rank, and has statewise full rank for $(\omega,\tilde{\omega})$ at regime R. Then for every p with $p(R) > 0$ and for every $\overline{k} > 0$ there is $\overline{K} > 0$ such that $k_i^p(T,\vec{\alpha}_{-i},\lambda_i,\delta,K) > \overline{k}$ for all (T,λ_i,δ,K) such that $\lambda_i^\omega \neq 0$, $\lambda_i^{\tilde{\omega}} \neq 0$, and $K > \overline{K}$. Therefore, if such $\vec{\alpha}_{-i}$ exists, then $k_i^p(T,\lambda_i) = \infty$ for all p and λ_i such that $p(R) > 0$, $\lambda_i^\omega \neq 0$ and $\lambda_i^{\tilde{\omega}} \neq 0$.*

Lemma 9. *Let λ_i be such that $\theta_{-i}^\omega \neq \theta_{-i}^{\tilde{\omega}}$ for any $(\omega,\tilde{\omega})$ satisfying $\omega \neq \tilde{\omega}$, $\lambda_i^\omega \neq 0$, and $\lambda_i^{\tilde{\omega}} \neq 0$. Let p be such that $k_i^p(T,e(\omega))$ is finite for all ω. Then $k_i^p(T,\lambda_i) \geq \sum_{\omega\in\Omega}|\lambda_i^\omega| k_i^p(T,e(\omega))$.*

Proof. Analogous to Lemma 6. *Q.E.D.*

Next we consider (T-LP) for single-state directions. Lemma 10 shows that under (Weak-CI), the scores of (T-LP) for single-state directions are bounded by the extreme values of belief-free review-strategy equilibrium payoffs of the known-state game.

Lemma 10. *Suppose that (Weak-CI) holds. Suppose also that the signal distribution has full support. Then $\liminf_{T\to\infty} k_i^p(T,\lambda_i) = N_i^{\omega,p}$ for $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega = 1$, and $\liminf_{T\to\infty} k_i^p(T,\lambda_i) = -n_i^{\omega,p}$ for $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega = -1$.*

Proof. We first consider direction λ_i such that $\lambda_i^\omega = 1$. Let $\vec{\alpha}_{-i}$ be such that for each R, player $-i$ chooses a pure action a_{-i}^R where a_{-i}^R is such that

$$\sum_{R\in\mathscr{R}} p(R) \min_{a_i \in R_i(\theta_i^\omega)} g_i^\omega(a_i, a_{-i}^R) = N_i^p.$$

Consider the problem (T-LP) for $(T,\vec{\alpha}_{-i},\lambda_i,\delta,K)$. Since (IFR) holds, $\vec{\alpha}_{-i}$ has individual full rank so that for each $\tilde{\omega} \neq \omega$, there is $x_i^{\tilde{\omega}}$ that makes player i indifferent in every period. Therefore we can ignore constraint (ii) for $\tilde{\omega} \neq \omega$. Section 3.3 of Yamamoto (2010) shows that under (Weak-CI), for any $\varepsilon > 0$ there is $\overline{T} > 0$ such that for any $T > \overline{T}$, there are $\overline{\delta} \in (0,1)$ and $K > 0$ such that for any $\delta \in (\overline{\delta},1)$, there is (v_i^ω, x_i^ω) such that $|v_i^\omega - N_i^{\omega,p}| < \varepsilon$ and all the remaining constraints of (T-LP) are satisfied. This shows that $\liminf_{T\to\infty} k_i^p(T,\lambda_i) \geq N_i^{\omega,p}$ for $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega = 1$. Also, it follows from Proposition 1 of Yamamoto (2010) that $k_i^p(T,\lambda_i) \leq N_i^{\omega,p}$ for any T. Therefore we have $\liminf_{T\to\infty} k_i^p(T,\lambda_i) = N_i^{\omega,p}$. A similar argument shows that $\liminf_{T\to\infty} k_i^p(T,\lambda_i) = -n_i^{\omega,p}$ for $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega = -1$. *Q.E.D.*

Combining the above three lemmas with Proposition 9, we obtain Proposition 10.

Example 4. *Secret Price-Cutting.* Here we apply Proposition 10 to a secret price-cutting game, and show that firms can maintain a self-enforcing cartel agreement even if they do now know how profitable the market is. Suppose that there are two firms and two possible states. In every period, firm i chooses either the high price $\overline{a}_i$ or the low price $\underline{a}_i$; that is, firm i's action set is $A_i = \{\overline{a}_i, \underline{a}_i\}$. Firm i's sales level $y_i \in Y_i$ is stochastically determined given firms' price $a \in A$. The joint distribution of $y = (y_1, y_2)$ is dependent on the state ω, and denoted by $\pi^\omega(\cdot|a)$. Firm i's utility is $u_i(a_i, y_i) = a_i y_i - c_i(y_i)$ where $c_i(y_i)$ is the production cost, and its expected payoff at state ω given a is $g_i^\omega = \sum_{y_i \in Y_i} \pi^\omega(y|a) u_i(a_i, y_i)$.

We assume that u_i and π are such that (SFR) and (Weak-CI) hold,[21] and such that the stage game is the prisoner's dilemma for both states; i.e., $(\overline{a}_1, \overline{a}_2)$ is efficient but $\underline{a}_i$ dominates $\overline{a}_i$ at each state. Then Proposition 10 applies so that for each $p \in \triangle\mathscr{R}$, the set $\times_{i\in I} \times_{\omega\in\Omega} [n_i^{\omega,p}, N_i^{\omega,p}]$ is in the limit set of sequential equilibrium payoffs as $\delta \to 1$. In particular for p such that $p(A) = 1$, we have $N_i^{\omega,p} = g_i^\omega(\overline{a}_i, \overline{a}_j)$ and $n_i^{\omega,p} = g_i^\omega(\underline{a}_i, \underline{a}_j)$ for each i and ω. Therefore the efficient payoff $g(\overline{a}_1, \overline{a}_2)$ can be approximated by a sequential equilibrium.

In this (nearly) efficient equilibrium, the infinite horizon is divided into T-period review phases, and in each review phase, a firm sets a constant price; that is, a firm chooses the high price for T periods or the low price for T periods. At the end of each review phase, a firm makes a statistical inference about its rival's price and about the true state using a sequence of past signals, and then in the next review phase, it chooses the high price or the low price contingently on that inference. The probability of setting the high price is judiciously chosen, and as a result, each firm's incentive compatibility is satisfied (specifically, a firm is indifferent between choosing the high price or the low price in each review phase) and near efficiency is attained state by state.

[21] Matsushima (2004) gives a condition under which the signal distribution of a secret price-cutting game is weakly conditionally independent.

Remark 5. In Lemma 10 (and hence in Proposition 10), we assume the signal distribution to be weakly conditionally independent. However, the result here is robust to a perturbation of the signal distribution; that is, $\liminf_{T\to\infty} k_i^p(T,\lambda_i)$ approximate $N_i^{\omega,p}$ for $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega = 1$, and approximate $-n_i^{\omega,p}$ for $\lambda_i \in \Lambda_i$ such that $\lambda_i^\omega = -1$, if the signal distribution is close to a weakly-conditionally-independent distribution. See Yamamoto (2010) for more details.

7 Limitations of BFXE

As argued in the introduction, BFXE have several nice robustness properties, specifically, BFXE are robust to any specification of a prior and robust to any specification of how players update their beliefs. However, BFXE are only a subset of sequential equilibria, and one may wonder if Pareto-efficiency can be improved by considering other sorts of equilibria. In this section, we provide an example where there is a sequential equilibrium Pareto-dominating all BFXE.

Suppose that there are two players and two states, so that $\boldsymbol{I} = \{1,2\}$ and $\Omega = \{\omega_1,\omega_2\}$. Assume that both players do not know the state, i.e., $\Theta_i = \{(\omega_1,\omega_2)\}$ for each i. In each stage game, player 1 chooses her action from the set $A_1 = \{UU,UD,DU,DD\}$, while player 2 chooses her action from the set $A_2 = \{L,M,R\}$. After players choose actions, player 1 observes the state perfectly and receive (possibly noisy) information about the actions. That is, the set of player 1's private signals is $\Sigma_1 = \Omega \times A$, and the marginal distribution $\pi_1^\omega(\cdot|a)$ is such that $\pi_1^\omega(\sigma_1|a) = 0$ for each $\sigma_1 = (\tilde{\omega},\tilde{a})$ such that $\omega \neq \tilde{\omega}$. On the other hand, player 2 observe the actions but does not observe the state; that is, the set of player 2's signals is $\Sigma_2 = A$, and the marginal distribution $\pi_2^\omega(\cdot|a)$ is such that $\pi_2^\omega(\sigma_2|a) = 0$ for each $\sigma_2 = (\tilde{a})$ such that $a \neq \tilde{a}$. The expected payoffs for state ω_1 is shown in the left panel, and those for state ω_2 is in the right.

	L	M	R
UU	1, 1	0, 0	0, 0
UD	0, 0	0, 1	0, 0
DU	0, 0	0, 0	0, 0
DD	0, 0	0, 0	0, 0

	L	M	R
UU	0, 1	0, 0	0, 0
UD	0, 0	1, 1	0, 0
DU	0, 0	0, 0	0, 0
DD	0, 0	0, 0	0, 0

Note that the efficient payoff is $((1,1),(1,1))$, and to approximate this payoff, players need to learn the state and condition their play on the state. (Specifically, players need to play (UU, L) at state ω_1 and play (UD, M) at state ω_2.) Since only player 1 can learn the state, information transmission from player 1 to player 2 is necessary for player 2 to condition her play on the state. The following strategy profile is a sequential equilibrium where player 1 transmits the information about the state to player 2:

- In period one, players will choose (DD, R).
- In period two, player 2 will choose R. Player 1 will choose DU if her signal in period one is $\sigma_1 = (\omega_1, a)$ for some a. Player 1 will choose DD if her signal in period one is $\sigma_1 = (\omega_2, a)$ for some a.
- In period three or later, player 1 will choose UU forever if her signal in period one is $\sigma_1 = (\omega_1, a)$ for some a and she chooses DU in period two. Player 1 will choose UD forever if her signal in period one is $\sigma_1 = (\omega_2, a)$ for some a and she chooses DD in period two. Otherwise, she will choose DD forever. Player 2 will choose L forever if player 1 chooses UD in period two. Player 2 will choose M forever if player 1 chooses DD in period two. Otherwise she will choose R forever.

Roughly speaking, the first period is used for "experiment" and then in the second period, player 1 tells the true state to player 2 through her action. After that players play the efficient action profile depending on the true state. It is easy to check that this strategy profile is a sequential equilibrium given any state ω, and approximate $((1,1),(1,1))$ when players are patient.

On the other hand, BFXE cannot approximate the efficient payoff $((1,1),(1,1))$. Indeed, for direction $\lambda_1 = (1,1)$, the score of (LP-Individual) is too low and the corresponding hyperplane $H_i^p(\lambda_1)$ does not contain $((1,1),(1,1))$.[22]

[22] Here we formally prove that BFXE cannot approximate $((1,1),(1,1))$. Consider (LP-Individual) for $\lambda_1 = (1,1)$. Given any p and $\vec{\alpha}_2$, it follows from constraint (i) that

$$\lambda_1 \cdot v_1 = \sum_{R \in \mathscr{R}} p(R) \sum_{a_2 \in A_2} \alpha_2^R(a_2) \left[\begin{array}{l} g_1^{\omega_1}(a_1^R, a_2) + g_1^{\omega_2}(a_1^R, a_2) \\ +\pi_2(a_1^R, a_2) \cdot (x_i^{\omega_1}(R, a_2) + x_i^{\omega_2}(R, a_2)) \end{array} \right]$$

for all $(a_1^R)_{R \in \mathscr{R}}$ such that $a_i^R \in R_i$ for all R. Then from constraint (iii), we have

$$\lambda_1 \cdot v_1 \leq \sum_{R \in \mathscr{R}} p(R) \sum_{a_2 \in A_2} \alpha_2^R(a_2) \left[g_1^{\omega_1}(a_1^R, a_2) + g_1^{\omega_2}(a_1^R, a_2) \right],$$

8 Concluding Remarks

In this paper, we study repeated games with private monitoring where players' payoffs and/or signal distributions are unknown. We look at a tractable subset of Nash equilibria, called BFXE, and characterize the limit set of equilibrium payoffs using a series of LP problems. If the individual and statewise full-rank conditions hold, then the limit equilibrium payoff set is isomorphic to the set of maps from states to belief-free equilibrium payoffs for the corresponding known-state game; that is, there are BFXE in which the payoffs are approximately the same as if players learn the true state and play a belief-free equilibrium for that state. Moreover, we combine BFXE with review strategies, and show that a larger payoff set can be achieved for games with weakly independent signal distribution.

Throughout this paper, we have assumed that players cannot communicate with each other. If we allow players to communicate at the end of every period, the folk theorem holds under some informational conditions. For example, suppose that there are three or more players and that they communicate after every stage game. Let $\pi^\omega(\cdot|a)$ be the distribution of a signal profile σ given an action profile a and a state ω, $\pi^\omega_{-i}(\cdot|a)$ be the marginal distribution of player $-i$'s signal σ_{-i}, and $\pi^\omega_{-ij}(\cdot|a)$ be the marginal distribution of player $-ij$'s signal σ_{-ij}. For each i, ω, and α, let $\Pi^{(i,\omega)}(\alpha)$ be a matrix with rows $\pi^\omega_{-i}(\cdot|a_i,\alpha_{-i})$ for all a_i. Likewise, for each i, j, ω, $\tilde{\omega}$, and α, let $\Pi^{(i,\omega)(j,\tilde{\omega})}(\alpha)$ be a matrix with rows $\pi^\omega_{-ij}(\cdot|a_i,\alpha_{-i})$ for all a_i and $\pi^{\tilde{\omega}}_{-ij}(\cdot|a_j,\alpha_{-j})$ for all a_j. For each (i,ω), a profile α has *individual full rank for* (i,ω) if the matrix $\Pi^{(i,\omega)}(\alpha)$ has rank $|A_i|$. Also, a profile α has individual full rank if it has individual full rank for all (i,ω). For (i,ω) and (j,ω) satisfying $i \neq j$, a profile α has *pairwise full rank for* (i,ω) *and* (j,ω) if $\Pi^{(i,\omega)(j,\omega)}(\alpha)$ has rank $|A_i|+|A_j|-1$. For (i,ω) and $(j,\tilde{\omega})$ satisfying $\omega \neq \tilde{\omega}$, a profile α has *statewise full rank for* (i,ω) *and* $(j,\tilde{\omega})$ if $\Pi^{(i,\omega)(j,\tilde{\omega})}(\alpha)$ has rank $|A_i|+|A_j|$. Then the folk theorem holds if the following conditions are satisfied:

- The set of feasible and individually rational payoffs is full dimensional.
- Every pure action profile α has individual full rank.

and the right-hand side is not greater than 1 for any p, $\vec{\alpha}_2$, and $(a_1^R)_{R\in\mathscr{R}}$. Therefore given any p, we have $k_1^p(\lambda_1) \leq 1$ so that $((1,1),(1,1))$ is not an element of the set Q^p.

- For each (i,ω) and (j,ω) satisfying $i \neq j$, there is α that has pairwise full rank for (i,ω) and (j,ω).

- For each (i,ω) and $(j,\tilde{\omega})$ satisfying $\omega \neq \tilde{\omega}$, there is α that has statewise full rank for (i,ω) and $(j,\tilde{\omega})$.

The formal proof is omitted, as it is very similar to Kandori and Matsushima (1998) and Fudenberg and Yamamoto (2010a).

As shown in Section 7, BFXE is only a subset of sequential equilibria, and a larger payoff set can be attained using "belief-based" equilibria. Unfortunately, belief-based equilibria do not have a recursive structure, and hence the the study of these equilibria would require different techniques. It remains open for future research.

Appendix

A.1 Proof of Proposition 3

Proposition 3. *For every* $\delta \in (0,1)$ *and* $p \in \triangle\mathscr{R}$, $E^p(\delta) = \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$.

Proof. First, we prove $E^p(\delta) \subseteq \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$. Let $v \in E^p(\delta)$, and let $s \in S$ be a BFXE yielding v. By definition of BFXE, in period one, no player wants to deviate from actions recommended by a public signal. Also, after every history $h^1 \in H^1$, the continuation strategy constitutes a BFXE with respect to p, so that the continuation payoff vector of $s|_{h^1}$ is in $E^p(\delta)$. Therefore, we conclude $v \in \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$, and hence $E^p(\delta) \subseteq \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$.

Next, we show $E^p(\delta) \supseteq \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$. Let $v \in \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$, and for each $i \in I$, let $\vec{\alpha}_{-i} \in \triangle\vec{A}_{-i}$ and $w_i : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to E_i^p(\delta)$ be such that v_i is individually ex-post generated by $(\vec{\alpha}_{-i}, w_i)$. For each $(R, a_{-i}, \sigma_{-i}) \in \mathscr{R} \times A_{-i} \times \Sigma_{-i}$, let $s_{-i}|_{(R,a_{-i},\sigma_{-i})} \in S_{-i}$ be player $-i$'s strategy such that for some $\tilde{s}_i \in S_i$, the profile $(\tilde{s}_i, s_{-i}|_{(R,a_{-i},\sigma_{-i})})$ is a BFXE with respect to p and gives the payoff $w_i^{R,\omega}(a_{-i}, \sigma_{-i})$ to player i for each state $\omega \in \Omega$. Let $s_{-i}^* \in S_{-i}$ be such that the action plan $\vec{\alpha}_{-i}$ is played in period one, and $s_{-i}|_{(R,a_{-i},\sigma_{-i})}$ is performed from period two on when player $-i$'s private history in period one is (R, a_{-i}, σ_{-i}). Consider the strategy profile $s^* = (s_1^*, s_2^*)$. From Proposition 1, the continuation play from period two induced by s^* is a BFXE with respect to p. Also, players do not want to deviate in period one for any state $\omega \in \Omega$, as v_i is individually ex-post generated by $(\vec{\alpha}_{-i}, w_i)$. Therefore, s^* is a BFXE with respect to p and attains v. This establishes $v \in E^p(\delta)$. *Q.E.D.*

A.2 Proof of Proposition 4

Proposition 4. *For each* $i \in I$, *let* W_i *be a subset of* $\mathbb{R}^{|\Omega|}$ *that is bounded and individually ex-post self-generating with respect to* (δ, p). *Then* $\times_{i\in I} W_i \subseteq E^p(\delta)$.

Proof. Let $v \in \times_{i\in I} W_i$, and we will construct a BFXE with payoff v. For each $i \in I$, player $-i$'s equilibrium strategy is determined in the following way. Since $v_i \in B_i^p(\delta, W_i)$, there are $\vec{\alpha}_{-i} \in \triangle\vec{A}_{-i}$ and $w_i : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to W_i$ such that v_i is individually ex-post generated by $(\vec{\alpha}_{-i}, w_i)$. Set player $-i$'s action plan in period one to be $\vec{\alpha}_{-i}$, i.e., $(s_{-i}(\theta_{-i}, h_{-i}^0, y^1))_{(\theta_{-i}, y^1)} = \vec{\alpha}_{-i}$. Also, for each one-period history $h^1_{-i} = (y^1, a^1_{-i}, \sigma^1_{-i})$, set $v_i|_{h^1_{-i}} = w_i(h^1_{-i}) \in W_i$. The play in later periods is determined recursively, using $v_i|_{h^t_{-i}}$ as a state variable. Given a $v_i|_{h^{t-1}_{-i}} \in W_i$ for $t \geq 2$, let $\vec{\alpha}_{-i} \in \triangle\vec{A}_{-i}$ and $w_i : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to W_i$ be such that $v_i|_{h^{t-1}_{-i}}$ is individually ex-post generated by $(\vec{\alpha}_{-i}, w_i)$, and then set $(s_{-i}(\theta_{-i}, h^{t-1}_{-i}, y^t))_{(\theta_{-i}, y^t)} = \vec{\alpha}_{-i}$ and $v_i|_{h^t_{-i}} = w_i^{y^t}(a^t_{-i}, \sigma^t_{-i}) \in W_i$ for $h^t_{-i} = (h^{t-1}_{-i}, (y^t, a^t_{-i}, \sigma^t_{-i}))$.

Note that this strategy always follows recommendation by public signals, that is, $s_i(\theta_i, h^{t-1}_i, R) \in R_i(\theta_i)$ for all $\theta_i \in \Theta_i$, $t \geq 1$, $h^{t-1}_i \in H^{t-1}_i$, and $R \in \mathscr{R}$. Also, because W_i is bounded and $\delta \in (0,1)$, payoffs are continuous at infinity. (See Fudenberg and Levine (1983) for details.) Thus finite approximations show that the specified strategy profile $s \in S$ generates v as an average payoff, and its continuation strategy $s|_{h^t}$ yields $v|_{h^t}$ for each $h^t \in H^t$. Moreover, it follows from (1) and (2) that at any moment of time and given any state $\omega \in \Omega$, a player is indifferent over all actions recommended by a realized public signal, and is worse off by deviating to other actions. Because payoffs are continuous at infinity, the one-shot deviation principle applies, and we conclude that s is a BFXE, as desired. *Q.E.D.*

A.3 Proof of Lemma 2

Lemma 2. *For every $\delta \in (0,1)$, $p \in \triangle\mathscr{R}$, and $i \in I$, $E^p_i(\delta) \subseteq coE^p_i(\delta) \subseteq Q^p_i$. Consequently, $E^p_i(\delta) \subseteq coE^p_i(\delta) \subseteq Q^p_i$.*

Proof. It is obvious that $E^p_i(\delta) \subseteq \text{co}E^p_i(\delta)$. Suppose $\text{co}E^p_i(\delta) \not\subseteq Q^p_i$. Then, since the score is a linear function, there are $v_i \in E^p_i(\delta)$ and λ_i such that $\lambda_i \cdot v_i > k^p_i(\lambda_i)$. In particular, since $E^p_i(\delta)$ is compact, there are $v^*_i \in E^p_i(\delta)$ and λ_i such that $\lambda_i \cdot v^*_i > k^p_i(\lambda_i)$ and $\lambda_i \cdot v^*_i \geq \lambda_i \cdot \tilde{v}_i$ for all $\tilde{v}_i \in \text{co}E^p_i(\delta)$. By definition, v^*_i is individually ex-post generated by w_i such that $w_i(R, a_{-i}, \sigma_{-i}) \in E^p_i(\delta) \subseteq \text{co}E^p_i(\delta) \subseteq H(\lambda_i, \lambda_i \cdot v^*_i)$ for all $\sigma_{-i} \in \Sigma_{-i}$. But this implies that $k^p_i(\lambda_i)$ is not the maximum score for direction λ_i, a contradiction. *Q.E.D.*

A.4 Proof of Lemma 4

Lemma 4. *For each $i \in I$, let W_i be a smooth subset of the interior of Q^p_i. Then there is $\overline{\delta} \in (0,1)$ such that for $\delta \in (\overline{\delta}, 1)$, $\times_{i \in I} W_i \subseteq E^p(\delta)$.*

Proof. From lemma 1(c), Q^p_i is bounded, and hence W_i is also bounded. Then, from Lemma 3, it suffices to show that W_i is locally ex-post generating, i.e., for each $v_i \in W_i$, there are $\delta_v \in (0,1)$ and an open neighborhood U_{v_i} of v_i such that $W \cap U_{v_i} \subseteq B(\delta_{v_i}, W)$.

First, consider v_i on the boundary of W_i. Let λ be normal to W_i at v_i, and let $k_i = \lambda_i \cdot v_i$. Since $W_i \subset Q_i \subseteq H_i^p(\lambda_i)$, there are $\vec{\alpha}_{-i}$, $\tilde{v}_i$, and $\tilde{w}_i$ such that $\lambda_i \cdot \tilde{v}_i > \lambda_i \cdot v_i = k_i$, $\tilde{v}_i$ is individually ex-post generated using $\vec{\alpha}_{-i}$ and $\tilde{w}_i$ for some $\tilde{\delta} \in (0,1)$, and $\tilde{w}_i(R, a_{-i}, \sigma_{-i}) \in H_i(\lambda_i, \lambda_i \cdot \tilde{v}_i)$ for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$, and $\sigma_{-i} \in \Sigma_{-i}$. For each $\delta \in (\tilde{\delta}, 1)$, let

$$w_i(R, a_{-i}, \sigma_{-i}) = \frac{\delta - \tilde{\delta}}{\delta(1-\tilde{\delta})} v_i + \frac{\tilde{\delta}(1-\delta)}{\delta(1-\tilde{\delta})} \left(\tilde{w}_i(R, a_{-i}, \sigma_{-i}) - \frac{v_i - \tilde{v}_i}{\tilde{\delta}} \right).$$

By construction, v_i is individually ex-post generated using $\vec{\alpha}_{-i}$ and w_i for δ, and there is $\kappa > 0$ such that $|w_i(R, a_{-i}, \sigma_{-i}) - v_i| < \kappa(1-\delta)$. Also, since $\lambda_i \cdot \tilde{v}_i > \lambda_i \cdot v_i = k_i$ and $\tilde{w}_i(R, a_{-i}, \sigma_{-i}) \in H_i(\lambda_i, \lambda_i \cdot \tilde{v}_i)$ for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$, and $\sigma_{-i} \in \Sigma_{-i}$, there is $\varepsilon > 0$ such that $\tilde{w}_i(R, a_{-i}, \sigma_{-i}) - \frac{v_i - \tilde{v}_i}{\tilde{\delta}}$ is in $H_i(\lambda_i, k_i - \varepsilon)$ for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$, and $\sigma_{-i} \in \Sigma_{-i}$. Then, $w_i(R, a_{-i}, \sigma_{-i}) \in H_i(\lambda_i, k_i - \frac{\tilde{\delta}(1-\delta)}{\delta(1-\tilde{\delta})}\varepsilon)$ for all $R \in \mathscr{R}$, $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$, and as in the proof of Theorem 3.1 of FL, it follows from the smoothness of W_i that $w_i(R, a_{-i}, \sigma_{-i}) \in \text{int}W_i$ for sufficiently large δ, i.e., v_i is individually ex-post generated with respect to intW_i using $\vec{\alpha}_{-i}$. To enforce u_i in the neighborhood of v_i, use this $\vec{\alpha}_{-i}$ and a translate of w_i.

Next, consider v_i in the interior of W_i. Choose λ_i arbitrarily, and let $\vec{\alpha}_{-i}$ and w_i be as in the above argument. By construction, v_i is individually ex-post generated by $\vec{\alpha}_{-i}$ and w_i. Also, $w_i(R, a_{-i}, \sigma_{-i}) \in \text{int}W_i$ for sufficiently large δ, since $|w_i(R, a_{-i}, \sigma_{-i}) - v_i| < \kappa(1-\delta)$ for some $\kappa > 0$ and $v_i \in \text{int}W_i$. Thus, v_i is enforced with respect to intW_i when δ is close to one. To enforce u_i in the neighborhood of v_i, use this $\vec{\alpha}_{-i}$ and a translate of w_i, as before. *Q.E.D.*

A.5 Proof of Lemma 5

Lemma 5. *Suppose that $\vec{\alpha}_{-i}$ has individual full rank, and has statewise full rank for $(\omega, \tilde{\omega})$ at regime R. Then for any p and λ_i satisfying $p(R) > 0$, $\lambda_i^\omega \neq 0$, and $\lambda_i^{\tilde{\omega}} \neq 0$. $k_i^p(\vec{\alpha}_{-i}, \lambda_i) = \infty$.*

Proof. First, we claim that for every $\bar{k} > 0$, there exist $(z_i^\omega(R, a_{-i}, \sigma_{-i}))_{(a_{-i}, \sigma_{-i})}$

and $(z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}))_{(a_{-i},\sigma_{-i})}$ such that

$$\sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^{\omega}}(a_{-i})\pi_{-i}^{\omega}(a)\cdot z_i^{\omega}(R,a_{-i}) = \frac{\overline{k}}{\delta p(R)\lambda_i^{\omega}} \tag{5}$$

for all $a_i \in A_i$,

$$\sum_{a_{-i\in A_{-i}}} \alpha_{-i}^{R,\theta_{-i}(\tilde{\omega})}(a_{-i})\pi_{-i}^{\tilde{\omega}}(a)\cdot z_i^{\tilde{\omega}}(R,a_{-i}) = 0 \tag{6}$$

for all $a_i \in A_i$, and

$$\lambda_i^{\omega} z_i^{\omega}(R,a_{-i},\sigma_{-i}) + \lambda_i^{\tilde{\omega}} z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}) = 0 \tag{7}$$

for all $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$, where $z_i^{\tilde{\omega}}(R,a_{-i}) = (z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}))_{\sigma_{-i}\in\Sigma_{-i}}$ and $z_i^{\tilde{\omega}}(R,a_{-i}) = (z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}))_{\sigma_{-i}\in\Sigma_{-i}}$. To prove that this system of equations indeed has a solution, eliminate (7) by solving for $z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i})$. Then, there remain $2|A_i|$ linear equations, and its coefficient matrix is $\Pi_{-i}^{(\omega,\tilde{\omega}),R}(\vec{\alpha}_{-i})$. Since statewise full rank implies that this coefficient matrix has rank $2|A_i|$, we can solve the system.

For each $\hat{R} \in \mathscr{R}$ and $\hat{\omega} \in \Omega$, let $(\tilde{w}_i^{\hat{\omega}}(\hat{R},a_{-i},\sigma_{-i}))_{(a_{-i},\sigma_{-i})}$ be such that

$$\sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{\hat{R},\theta_{-i}(\hat{\omega})}(a_{-i})\left[(1-\delta)g_i^{\hat{\omega}}(a) + \delta\pi_{-i}^{\hat{\omega}}(a)\cdot\tilde{w}_i^{\hat{\omega}}(\hat{R},a_{-i})\right] = 0 \tag{8}$$

for all $a_i \in A_i$. In words, the continuation payoffs $\tilde{w}_i$ are chosen so that for each state $\hat{\omega}$ and for each realized public signal $\hat{R}$, player i is indifferent among all actions and his overall payoff is zero. Note that this system has a solution, since α has individual full rank.

Let $\overline{k} > \max_{(\hat{R},a_{-i},\sigma_{-i})} \lambda_i \cdot \tilde{w}_i(\hat{R},a_{-i},\sigma_{-i})$, and choose $(z_i^{\omega}(R,a_{-i},\sigma_{-i}))_{(a_{-i},\sigma_{-i})}$ and $(z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}))_{(a_{-i},\sigma_{-i})}$ to satisfy (5) through (7). Then, let

$$w_i^{\hat{\omega}}(\hat{R},a_{-i},\sigma_{-i}) = \begin{cases} \tilde{w}_i^{\omega}(R,a_{-i},\sigma_{-i}) + z_i^{\omega}(R,a_{-i},\sigma_{-i}) & \text{if} \quad (\hat{R},\hat{\omega}) = (R,\omega) \\ \tilde{w}_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}) + z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}) & \text{if} \quad (\hat{R},\hat{\omega}) = (R,\tilde{\omega}) \\ \tilde{w}_i^{\hat{\omega}}(\hat{R},a_{-i},\sigma_{-i}) & \text{otherwise} \end{cases}$$

for each $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$. Also, let

$$v_i^{\hat{\omega}} = \begin{cases} \frac{\overline{k}}{\lambda_i^{\omega}} & \text{if} \quad \hat{\omega} = \omega \\ 0 & \text{otherwise} \end{cases} .$$

We claim that this (v_i, w_i) satisfies constraints (i) through (iii) in the LP problem. It follows from (8) that constraints (i) and (ii) are satisfied for all $\hat{\omega} \neq \omega, \tilde{\omega}$. Also, using (5) and (8), we obtain

$$\begin{aligned}
&\sum_{\hat{R}\in\mathscr{R}} p(\hat{R}) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{\hat{R},\theta_{-i}^{\omega}}(a_{-i}) \left[(1-\delta)g_i^{\omega}(a_i,a_{-i}) + \delta\pi_{-i}^{\hat{\omega}}(a)\cdot w_i^{\omega}(\hat{R},a_{-i})\right] \\
=&\sum_{\hat{R}\in\mathscr{R}} p(\hat{R}) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{\hat{R},\theta_{-i}^{\omega}}(a_{-i}) \left[(1-\delta)g_i^{\omega}(a_i,a_{-i}) + \delta\pi_{-i}^{\hat{\omega}}(a)\cdot \tilde{w}_i^{\omega}(\hat{R},a_{-i})\right] \\
&+\delta p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^{\omega}}(a_{-i})\pi_{-i}^{\hat{\omega}}(a)\cdot z_i^{\omega}(R,a_{-i}) \\
=&\frac{\overline{k}}{\lambda_i^{\omega}}
\end{aligned}$$

for all $a_i \in A_i$. This shows that (v_i, w_i) satisfies constraints (i) and (ii) for ω. Likewise, from (6) and (8), (v_i, w_i) satisfies constraints (i) and (ii) for $\tilde{\omega}$. Furthermore, using (7) and $\overline{k} > \max_{(\hat{R},a_{-i},\sigma_{-i})} \lambda_i \cdot \tilde{w}_i(\hat{R}, a_{-i}, \sigma_{-i})$, we have

$$\begin{aligned}
&\lambda_i \cdot w_i(R,a_{-i},\sigma_{-i}) \\
=&\lambda_i \cdot \tilde{w}_i(R,a_{-i},\sigma_{-i}) + \lambda_i^{\omega} z_i^{\omega}(R,a_{-i},\sigma_{-i}) + \lambda_i^{\tilde{\omega}} z_i^{\tilde{\omega}}(R,a_{-i},\sigma_{-i}) \\
=&\lambda_i \cdot \tilde{w}_i(R,a_{-i},\sigma_{-i}) \\
<&\overline{k} = \lambda_i \cdot v_i
\end{aligned}$$

for all $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$, and we have

$$\lambda_i \cdot w_i(\hat{R},a_{-i},\sigma_{-i}) = \lambda_i \cdot \tilde{w}_i(\hat{R},a_{-i},\sigma_{-i}) < \overline{k} = \lambda_i \cdot v_i$$

for all $\hat{R} \neq R$, $a_{-i} \in A_{-i}$, and $\sigma_{-i} \in \Sigma_{-i}$. Hence, constraint (iii) holds.

Therefore, $k_i^p(\vec{\alpha}_{-i}, \lambda_i) \geq \lambda_i \cdot v_i = \overline{k}$. Since $\overline{k}$ can be arbitrarily large, we conclude $k_i^p(\vec{\alpha}_{-i}, \lambda_i) = \infty$. *Q.E.D.*

A.6 Proof of Lemma 6

Lemma 6. *Suppose that (IFR) holds. Let λ_i be such that $\theta_{-i}^{\omega} \neq \theta_{-i}^{\tilde{\omega}}$ for any $(\omega, \tilde{\omega})$ satisfying $\omega \neq \tilde{\omega}$, $\lambda_i^{\omega} \neq 0$, and $\lambda_i^{\tilde{\omega}} \neq 0$. Then $k_i^p(\lambda_i) \geq \sum_{\omega\in\Omega} |\lambda_i^{\omega}| k_i^p(e(\omega))$, and as a result $\bigcap_{\omega\in\Omega} H_i^p(e(\omega)) \subseteq H_i^p(\lambda_i)$.*

Proof. If $k_i^p(\lambda_i) = \infty$, then the result is obvious. So assume $k_i^p(\lambda_i) < \infty$. First, we claim

$$k_i^p(\vec{\alpha}_{-i}, \lambda_i) \geq \sum_{\omega \in \Omega} |\lambda_i^\omega| k_i^p(\vec{\alpha}_{-i}, e(\omega)) \tag{9}$$

for each $\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}$. To show this, consider the LP problem for $(\vec{\alpha}_{-i}, \lambda_i)$ but constraint (iii) is replaced with a more restrictive condition

(iii′) $\lambda_i^\omega v_i(\omega) \geq \lambda_i^\omega w_i^\omega(R, a_{-i}, \sigma_{-i})$ for all $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$.

Let $\tilde{k}_i^p(\vec{\alpha}_{-i}, \lambda_i)$ denote the solution to this new problem. Since condition (iii′) does not allow utility transfer across different states, considering this new LP problem is equivalent to solving a separate LP problem for each state $\omega \in \Omega$ in isolation. Thus we have $\tilde{k}_i^p(\vec{\alpha}_{-i}, \lambda_i) = \sum_{\omega \in \Omega} |\lambda_i^\omega| k_i^p(\vec{\alpha}_{-i}, e(\omega))$. (Here we use the fact that $\tilde{k}_i^p(\vec{\alpha}_{-i}, \lambda_i)$ and $k_i^p(\vec{\alpha}_{-i}, e(\omega))$ are finite under (IFR).) Since $k_i^p(\vec{\alpha}_{-i}, \lambda_i) \geq \tilde{k}_i^p(\vec{\alpha}_{-i}, \lambda_i)$, (9) follows.

Recall that $e(\omega)$ has only one nonzero component $e^\omega(\omega)$. Hence, the maximal score $k_i^p(\vec{\alpha}_{-i}, e(\omega))$ for direction $e(\omega)$ depends on $\vec{\alpha}_{-i}$ only through $\alpha_{-i}(\omega) = (\alpha_{-i}^{R, \theta_{-i}^\omega})_{R \in \mathscr{R}}$; so we write it by $k_i^p(\alpha_{-i}(\omega), e(\omega))$ to emphasize the dependence. Under (IFR), we have $k_i^p(\vec{\alpha}_{-i}, e(\omega)) > -\infty$ for each $\vec{\alpha}_{-i}$ and ω. Therefore we have

$$\begin{aligned} \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} \sum_{\omega \in \Omega} |\lambda_i^\omega| k_i^p(\vec{\alpha}_{-i}, e(\omega)) &= \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} \sum_{\omega \in \Omega} |\lambda_i^\omega| k_i^p(\alpha_{-i}^{\theta_{-i}^\omega}, e(\omega)) \\ &= \sum_{\omega \in \Omega} \sup_{\alpha_{-i}(\omega)} |\lambda_i^\omega| k_i^p(\alpha_{-i}(\omega), e(\omega)) \\ &= \sum_{\omega \in \Omega} \sup_{\vec{\alpha}_{-i} \in \triangle \vec{A}_{-i}} |\lambda_i^\omega| k_i^p(\vec{\alpha}_{-i}, e(\omega)) \\ &= \sum_{\omega \in \Omega} |\lambda_i^\omega| k_i^p(e(\omega)). \end{aligned}$$

Here the second equality comes from the fact that $\theta_{-i}^\omega \neq \theta_{-i}^{\tilde{\omega}}$ for any pair $(\omega, \tilde{\omega})$ such that $\omega \neq \tilde{\omega}$. Plugging this into (9), we obtain the desired result. *Q.E.D.*

A.7 Proof of Proposition 9

Proposition 9. *Suppose that the signal distribution has full support. Let T and p be such that* $\dim Q_i^p(T) = |\Omega|$ *for each $i \in I$. Then the set $\times_{i \in I} Q_i^p(T)$ is in the limit set of sequential equilibrium payoffs as $\delta \to 1$.*

To prove this proposition, we begin with some preliminary results.

Definition 11. Player i's payoff $v_i = (v_i^\omega)_{\omega\in\Omega} \in \mathbb{R}^{|\Omega|}$ is *individually ex-post generated with respect to* (T,δ,W_i,p) if there is an action plan $\vec{\alpha}_{-i} \in \triangle\vec{A}_{-i}$ and a function $w_i : \mathscr{R} \times A_{-i} \times (\Sigma_{-i})^T \to W_i$ such that

$$v_i^\omega = \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} (1-\delta^T)g_i^\omega(a_i^R,a_{-i}) \\ +\delta^T \pi_{-i}^{T,\omega}(a_i^R,a_{-i}) \cdot w_i^{T,\omega}(R,a_{-i}) \end{array} \right]$$

for all $\omega\in\Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ satisfying $a_i^R \in R_i(\theta_i^\omega)$ for each $R\in\mathscr{R}$, and

$$v_i^\omega \geq \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R,\theta_{-i}^\omega}(a_{-i}) \left[\begin{array}{l} (1-\delta^T)g_i^{T,\omega}(s_i^{T,R},a_{-i}) \\ +\delta^T \pi_{-i}^{T,\omega}(s_i^{T,R},a_{-i}) \cdot w_i^{T,\omega}(R,a_{-i}) \end{array} \right]$$

for all $\omega\in\Omega$ and $(s_i^{T,R})_{R\in\mathscr{R}}$ satisfying $s_i^{T,R} \in S_i^T$ for each $R\in\mathscr{R}$.

Let $B_i^p(T,\delta,W_i)$ be the set of all v_i individually ex-post generated with respect to (T,δ,W_i,p). A subset W_i of $\mathbb{R}^{|\Omega|}$ is individually ex-post self-generating with respect to (T,δ,p) if $W_i \subseteq B_i^p(T,\delta,W_i,p)$

Proposition 10. *For each $i\in I$, let W_i be a subset of $\mathbb{R}^{|\Omega|}$ that is bounded and individually ex-post self-generating with respect to (T,δ,p). Then $\times_{i\in I}W_i$ is in the set of sequential equilibrium payoffs with public randomization p for δ.*

Proof. Analogous to Proposition 4. *Q.E.D.*

Given any $v_i \in \mathbb{R}^{|\Omega|}$, $\lambda_i \in \Lambda_i$, $\varepsilon > 0$, $K > 0$, and $\delta \in (0,1)$, let $G_{v_i,\lambda_i,\varepsilon,K,\delta}$ be the set of all $v_i' \in \mathbb{R}^{|\Omega|}$ such that $\lambda_i \cdot v_i \geq \lambda_i \cdot v_i' + (1-\delta)\varepsilon$ and such that v_i' is within $(1-\delta)K$ of v_i. (See Figure 1, where this set is labeled "G.")

Definition 12. A subset W_i of $\mathbb{R}^{|\Omega|}$ is *uniformly decomposable with respect to* (T,p) if there are $\varepsilon > 0$, $K > 0$, and $\overline{\delta} \in (0,1)$ such that for any $v_i \in W_i$, $\delta \in (\overline{\delta},1)$, and $\lambda_i \in \Lambda_i$, there are $\vec{\alpha}_{-i}$ and $w_i : \mathscr{R} \times A_{-i} \times (\Sigma_{-i})^T \to W_i$ such that $(\vec{\alpha}_{-i}, v_i)$ is enforced by w_i for δ and such that $w_i(R,a_{-i},\sigma_{-i}^1,\cdots,\sigma_{-i}^T) \in G_{v_i,\lambda_i,\varepsilon,K,\delta^T}$ for all $(R,a_{-i},\sigma_{-i}^1,\cdots,\sigma_{-i}^T)$.

Lemma 7. *Suppose that a subset W_i of $\mathbb{R}^{|\Omega|}$ is smooth, bounded, and uniformly decomposable with respect to (T,p). Then there is $\overline{\delta} \in (0,1)$ such that W_i is individually ex-post self-generating with respect to (T,δ,p) for any $\delta \in (\overline{\delta},1)$.*

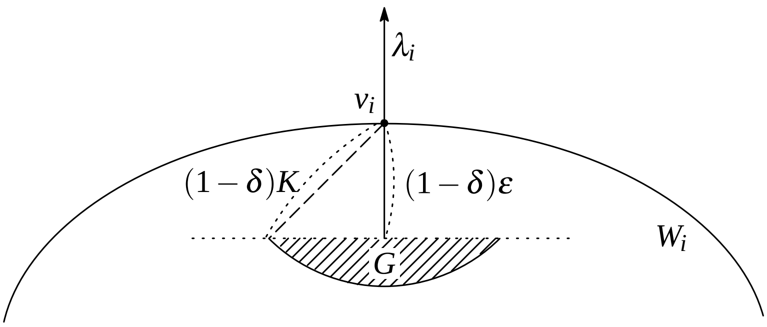

Figure 1: Set G.

Proof. Analogous to Fudenberg and Yamamoto (2010c). *Q.E.D.*

Lemma 8. *Any smooth subset W_i of the interior of $Q_i^p(T)$ is bounded and uniformly decomposable with respect to (T,p).*

Proof. As in Lemma 1, one can check that $Q_i^p(T)$ is bounded, and so is W_i. Let $\tilde{\varepsilon} > 0$ be such that $|v_i' - v_i''| > \tilde{\varepsilon}$ for all $v_i' \in W_i$ and $v_i'' \in Q_i^p(T)$. By definition, for every $\lambda_i \in \Lambda_i$, $k_i^p(T,\lambda_i) > \max_{v_i' \in W_i} \lambda_i \cdot v_i' + \tilde{\varepsilon}$. Therefore for each $\lambda_i \in \Lambda_i$, there are $\overline{\delta}_{\lambda_i} \in (0,1)$ and $K_{\lambda_i} > 0$ such that for any $\delta \in (\overline{\delta}_{\lambda_i}, 1)$, there is $\vec{\alpha}_{-i,\lambda_i,\delta}$ such that $k_i^p(T, \vec{\alpha}_{-i,\lambda_i,\delta}, \lambda_i, \delta, K_{\lambda_i}) > \max_{v_i' \in W_i} \lambda_i \cdot v_i' + \tilde{\varepsilon}$.

Given λ_i and $\delta \in (\overline{\delta}_{\lambda_i}, 1)$, let $\tilde{v}_{i,\lambda_i,\delta} \in \mathbb{R}^{|\Omega|}$ and $x_{i,\lambda_i,\delta} : \mathscr{R} \times A_{-i} \times (\Sigma_{-i})^T \to \mathbb{R}^{|\Omega|}$ be such that all the constraints of the LP problem for $(T, \vec{\alpha}_{-i,\lambda_i,\delta}, \lambda_i, \delta, K_{\lambda_i})$ are satisfied and such that $\lambda_i \cdot \tilde{v}_{i,\lambda_i,\delta} > \max_{v_i' \in W_i} \lambda_i \cdot v_i' + \tilde{\varepsilon}$. Then for each $v_i \in W_i$, let $w_{i,\lambda_i,\delta,v_i} : \mathscr{R} \times A_{-i} \times (\Sigma_{-i})^T \to \mathbb{R}^{|\Omega|}$ be such that

$$w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) = v_i + \frac{1-\delta^T}{\delta^T}(v_i - \tilde{v}_{i,\lambda_i,\delta} + x_{i,\lambda_i,\delta}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T))$$

for each $(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)$. By construction, $(\vec{\alpha}_{-i,\lambda_i,\delta}, v_i)$ is enforced by $w_{i,\lambda_i,\delta,v_i}$ for δ. Also, letting $\varepsilon = \frac{\tilde{\varepsilon}}{2}$ and $\tilde{K}_{\lambda_i} = K_{\lambda_i} + \sup_{v_i' \in W_i} \sup_{\delta \in (\overline{\delta}_{\lambda_i}, 1)} |v_i' - \tilde{v}_{i,\lambda_i,\delta}|$, it follows that $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) \in G_{v_i,\lambda_i,2\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$. (To see this, note first that the pair $(\tilde{v}_{i,\lambda_i,\delta}, x_{i,\lambda_i,\delta})$ satisfies constraints (i) and (iv) of the LP problem so that $\sup_{\delta \in (\overline{\delta}_{\lambda_i}, 1)} |\tilde{v}_{i,\lambda_i,\delta}| \leq \max_{a \in A} |(g_i^\omega(a))_{\omega \in \Omega}| + K_{\lambda_i}$. This and the boundedness of W_i show that $\tilde{K}_{\lambda_i} < \infty$. Since $\lambda_i \cdot x_{i,\lambda_i,\delta}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) \leq 0$ and $\lambda_i \cdot \tilde{v}_{i,\lambda_i,\delta} > \max_{v_i' \in W_i} \lambda_i \cdot v_i' + \tilde{\varepsilon} \geq \lambda_i \cdot v_i + \tilde{\varepsilon}$, it follows that $\lambda_i \cdot w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) \leq \lambda_i \cdot v_i - \frac{1-\delta^T}{\delta^T}\tilde{\varepsilon} < \lambda_i \cdot v_i - (1-\delta^T)\tilde{\varepsilon}$. Also, $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)$ is within $\frac{1-\delta^T}{\delta^T}\tilde{K}_{\lambda_i}$ of v_i, as $|x_{i,\lambda_i,\delta}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)| < K_{\lambda_i}$.)

Note that for each $\lambda_i \in \Lambda_i$, there is an open set $U_{\lambda_i,\delta} \subseteq \mathbb{R}^{|\Omega|}$ containing λ_i such that $G_{v_i,\lambda_i,2\varepsilon,\tilde{K}_{\lambda_i},\delta^T} \subseteq G_{v_i,\lambda_i',\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$ for any $v_i \in W_i$, $(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)$, and $\lambda_i' \in \Lambda_i \cap U_{\lambda_i,\delta,v_i}$. (See Figure 2, where $G_{v_i,\lambda_i,2\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$ and $G_{v_i,\lambda_i',\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$ are labeled "G" and "G'," respectively.) Then we have $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) \in G_{v_i,\lambda_i',\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$ for any $v_i \in W_i$, $(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)$, and $\lambda_i' \in \Lambda_i \cap U_{\lambda_i,\delta,v_i}$, since $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) \in G_{v_i,\lambda_i,2\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$.

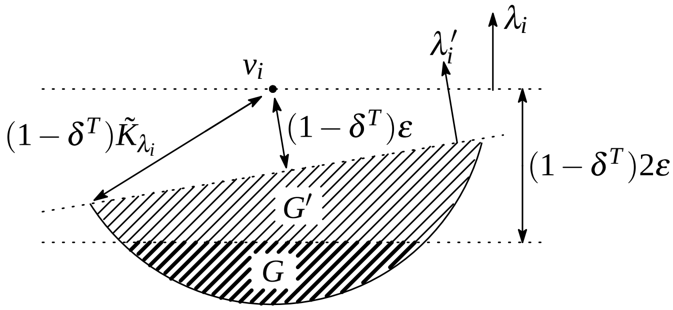

Figure 2: $G \subseteq G'$.

The set Λ_i is compact, so $\{U_{\lambda_i,\delta}\}_{\lambda_i \in \Lambda_i}$ has a finite subcover $\{U_{\lambda_i,\delta}\}_{\lambda_i \in \tilde{\Lambda}}$. For each v_i and λ_i, let $\vec{\alpha}^*_{-i,\lambda_i,\delta} = \vec{\alpha}_{-i,\lambda_i',\delta}$ and $w^*_{i,\lambda_i,\delta,v_i} = w_{i,\lambda_i',\delta,v_i}$, where $\lambda_i' \in \tilde{\Lambda}_i$ is such that $\lambda_i \in U_{\lambda_i',\delta}$. Let $K = \max_{\lambda_i \in \tilde{\Lambda}_i} \tilde{K}_{\lambda_i}$. Then $(\vec{\alpha}^*_{-i,\lambda_i,\delta}, v_i)$ is enforced by $w^*_{i,\lambda_i,\delta,v_i}$ and $w^*_{i,\lambda_i,\delta,v_i}$ chooses the continuation payoffs from the set $G_{v_i,\lambda_i,\varepsilon,K,\delta^T}$. Note that now K is independent of λ_i, and thus the proof is completed. *Q.E.D.*

From the above lemmas, Proposition 9 follows.

A.8 Proof of Lemma 8

Lemma 8. *Suppose that (IFR) holds. Suppose also that $\vec{\alpha}_{-i}$ has individual full rank, and has statewise full rank for $(\omega, \tilde{\omega})$ at regime R. Then for every p with $p(R) > 0$ and for every $\overline{k} > 0$ there is $\overline{K} > 0$ such that $k_i^p(T, \vec{\alpha}_{-i}, \lambda_i, \delta, K) > \overline{k}$ for all $(T, \lambda_i, \delta, K)$ such that $\lambda_i^\omega \neq 0$, $\lambda_i^{\tilde{\omega}} \neq 0$, and $K > \overline{K}$. Therefore, if such $\vec{\alpha}_{-i}$ exists, then $k_i^p(T, \lambda_i) = \infty$ for all p and λ_i such that $p(R) > 0$, $\lambda_i^\omega \neq 0$ and $\lambda_i^{\tilde{\omega}} \neq 0$.*

Proof. Since (IFR) holds, there is $z_i : A_{-i} \times \Sigma_{-i} \to \mathbb{R}^{|\Omega|}$ such that

$$g_i^{\tilde{\omega}}(a) + \sum_{\sigma_{-i} \in \Sigma_{-i}} \pi_{-i}^{\tilde{\omega}}(\sigma_{-i}|a) z_i^{\tilde{\omega}}(a_{-i}, \sigma_{-i}) = g_i^{\tilde{\omega}}(a_i', a_{-i}) + \sum_{\sigma_{-i} \in \Sigma_{-i}} \pi_{-i}^{\tilde{\omega}}(\sigma_{-i}|a_i', a_{-i}) z_i^{\tilde{\omega}}(a_{-i}, \sigma_{-i})$$

for all $\tilde{\omega} \in \Omega$, $a \in A$, and $a_i' \neq a_i$. That is, z_i is chosen in such a way that player i is indifferent over all actions in a one-shot game if she receives a payment $z_i(a_{-i}, \sigma_{-i})$ after play. In particular we can choose z_i so that

$$\lambda_i \cdot z_i(a_{-i}, \sigma_{-i}) \leq 0$$

for all $a_{-i} \in A_{-i}$ and $\sigma_{-i} \in \Sigma_{-i}$. Let $\hat{v}_i \in \mathbb{R}^{|\Omega|}$ be player i's payoff of the one-shot game with payment z_i when player $-i$ plays $\vec{\alpha}_{-i}$ and a public signal R follows a distribution p; that is,

$$\hat{v}_i^{\omega} = \sum_{R \in \mathscr{R}} p(R) \sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R, \theta_{-i}^{\omega}}(a_{-i}) \left[g_i^{\omega}(a, \delta) + \sum_{\sigma_{-i} \in \Sigma_{-i}} \pi_{-i}^{\omega}(\sigma_{-i}|a) z_i^{\omega}(a_{-i}, \sigma_{-i}) \right]$$

for some a_i.

Also, it follows from Lemma 5 that for every $\overline{k} > 0$, there are $\tilde{v}_i \in \mathbb{R}^{|\Omega|}$ and $\tilde{x}_i : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to \mathbb{R}^{|\Omega|}$ such that $(\tilde{v}_i, \tilde{x}_i)$ satisfies constraints (i) through (iii) of (LP-Individual) and such that $\lambda_i \cdot \tilde{v}_i \geq T\overline{k} + (T-1)|\lambda_i \cdot \hat{v}_i|$. Let

$$v_i = \frac{1-\delta}{1-\delta^T} \left(\tilde{v}_i + \sum_{\tau=2}^{T} \delta^{\tau-1} \hat{v}_i \right)$$

and

$$x_i(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T) = \frac{1-\delta}{1-\delta^T} \left(\tilde{x}_i(R, a_{-i}, \sigma_{-i}^1) + \sum_{\tau=2}^{T} \delta^{\tau-1} z_i(a_{-i}, \sigma_{-i}^{\tau}) \right).$$

Then this (v_i, x_i) satisfies constraints (i) through (iii) of (T-LP). Also, letting

$$K > \max_{(R, a_{-i}, \sigma_{-i})} |\tilde{x}_i(R, a_{-i}, \sigma_{-i})| + \max_{(a_{-i}, \sigma_{-i})} (T-1)|z_i(a_{-i}, \sigma_{-i})|,$$

condition (iv) also holds. Since $\lambda_i \cdot v_i \geq \overline{k}$, the lemma follows. *Q.E.D.*

References

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58, 1041-1063.

Aumann, R., and M. Maschler (1995): *Repeated Games with Incomplete Information*. MIT Press, Cambridge, MA. With the collaboration of R.E. Stearns.

Bhaskar, V., G.J. Mailath, and S. Morris (2008): "Purification in the Infinitely Repeated Prisoner's Dilemma," *Review of Economic Dynamics* 11, 515-528.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory* 102, 40-69.

Chen, B. (2010): "A Belief-Based Approach to the Repeated Prisoners' Dilemma with Asymmetric Private Monitoring," *Journal of Economic Theory* 145, 402-420.

Cripps, M., J. Ely, G. Mailath, and L. Samuelson (2008): "Common Learning," *Econometrica* 76, 909-933.

Cripps, M., and J. Thomas (2003): "Some Asymptotic Results in Discounted Repeated Games of One-Side Incomplete Information," *Mathematics of Operations Research* 28, 433-462.

Ely, J., J. Hörner, and W. Olszewski (2005): "Belief-Free Equilibria in Repeated Games," *Econometrica* 73, 377-415.

Ely, J., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory* 102, 84-105.

Fong, K., O. Gossner, J. Hörner, and Y. Sannikov (2010): "Efficiency in a Repeated Prisoner's Dilemma with Imperfect Private Monitoring," mimeo.

Forges, F. (1984): "Note on Nash Equilibria in Infinitely Repeated Games with Incomplete Information," *International Journal of Game Theory* 13, 179-187.

Fuchs, W. (2007): "Contracting with Repeated Moral Hazard and Private Evaluations," *American Economic Review* 97, 1432-1448.

Fudenberg, D., and D.K. Levine (1983): "Subgame-Perfect Equilibria of Finite and Infinite Horizon Games," *Journal of Economic Theory* 31, 251-268.

Fudenberg, D., and D.K. Levine (1991): "Approximate Equilibria in Repeated Games with Imperfect Private Information," *Journal of Economic Theory* 54, 26-47.

Fudenberg, D., and D.K. Levine (1994): "Efficiency and Observability in Games with Long-Run and Short-Run Players," *Journal of Economic Theory* 62, 103-135.

Fudenberg, D., D.K. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica* 62, 997-1040.

Fudenberg, D., D.K. Levine, and S. Takahashi (2007): "Perfect Public Equilibrium when Players are Patient," *Games and Economic Behavior* 61, 27-49.

Fudenberg, D., and Y. Yamamoto (2010a): "Repeated Games where the Payoffs and Monitoring Structure are Unknown," *Econometrica* 78, 1673-1710.

Fudenberg, D., and Y. Yamamoto (2010b): "Learning from Private Information in Noisy Repeated Games," mimeo.

Fudenberg, D., and Y. Yamamoto (2010c): "The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring," mimeo.

Gossner, O., and N. Vieille (2003): "Strategic Learning in Games with Symmetric Information," *Games and Economic Behavior* 42, 25-47.

Harrington, J., and A. Skrzypacz (2010): "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices," forthcoming in *American Economic Review*.

Hart, S. (1985): "Nonzero-Sum Two-Person Repeated Games with Incomplete Information," *Mathematics of Operations Research* 10, 117-153.

Hörner, J., and S. Lovo (2009): "Belief-Free Equilibria in Games with Incomplete Information," *Econometrica* 77, 453-487.

Hörner, J., S. Lovo, and T. Tomala (2009): "Belief-Free Equilibria in Games with Incomplete Information: Characterization and Existence," mimeo.

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

Hörner, J., and W. Olszewski (2009): "How Robust is the Folk Theorem with Imperfect Public Monitoring?," *Quarterly Journal of Economics* 124, 1773-1814.

Ishii, Y. (2009): “Folk Theorem with a Continuum of Public Signals,” mimeo.

Kandori, M. (2002): “Introduction to Repeated Games with Private Monitoring,” *Journal of Economic Theory* 102, 1-15.

Kandori, M. (2010): “Weakly Belief-Free Equilibria in Repeated Games with Private Monitoring,” forthcoming in *Econometrica*.

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

Kandori, M., and I. Obara (2006): “Efficiency in Repeated Games Revisited: the Role of Private Strategies,” *Econometrica* 72, 499-519.

Lehrer, E. (1990): “Nash Equilibria of n-Player Repeated Games with Semi-Standard Information,” *International Journal of Game Theory* 19, 191-217.

Levin, J. (2003): “Relational Incentive Contracts,” *American Economic Review* 93, 835-857.

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

Mailath, G.J., and S. Morris (2006): “Coordination Failure in Repeated Games with Almost-Public Monitoring,” *Theoretical Economics* 1, 311-340.

Mailath, G.J., and W. Olszewski (2010): “Folk Theorems with Bounded Recall and (Almost) Perfect Monitoring,” forthcoming in *Games and Economic Behavior*.

Mailath, G.J., and L. Samuelson (2006): *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, New York, NY.

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

Miller, D. (2009): “Optimal Ex Post Incentive Compatible Equilibria in Repeated Games of Private Information,” mimeo.

Olszewski, W. (2007): “A Simple Exposition of Belief-Free Equilibria in Repeated Games,” *Economics Bulletin* 58, 1-16.

Park, J-H. (2010): “Enforcing International Trade Agreements with Imperfect Private Monitoring,” mimeo.

Pecorino, P. (1998): "Is There a Free-Rider Problem in Lobbying? Endogenous Tariffs, Trigger Strategies, and the Number of Firms," *American Economic Review* 88, 652-660.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory* 102, 70-83.

Radner, R. (1985): "Repeated Principal-Agent Games with Discounting," *Econometrica* 53, 1173-1198.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory* 76, 345-361.

Sorin, S. (1984): "Big Match with Lack of Information on One Side (Part I)," *International Journal of Game Theory* 13, 201-255.

Sorin, S. (1985): "Big Match with Lack of Information on One Side (Part II)," *International Journal of Game Theory* 14, 173-204.

Stigler, G.J. (1964): "A Theory of Oligopoly," *Journal of Political Economy* 72, 44-61.

Sugaya, T. (2009): "Folk Theorem in a Prisoners' Dilemma without Conditional Independence," mimeo.

Sugaya, T. (2010a): "Belief-Free Review-Strategy Equilibrium without Conditional Independence," mimeo.

Sugaya, T. (2010b): "Folk Theorem in a Two-Player Prisoners' Dilemma," mimeo.

Sugaya, T., and S. Takahashi (2010): "Coordination Failure in Repeated Games with Private Monitoring," mimeo.

Takahashi, S. (2010): "Community Enforcement when Players Observe Partners' Past Play," *Journal of Economic Theory* 145, 42-62.

Telser, L. (1960): "Why Should Manufacturers Want Fair Trade?," *Journal of Law and Economics* 3, 86-105.

Varian, H. R. (1980): "A Model of Sales," *American Economic Review* 70, 651-659.

Wiseman, T. (2005): "A Partial Folk Theorem for Games with Unknown Payoff Distributions," *Econometrica* 73, 629-645.

Wiseman, T. (2010) "A Partial Folk Theorem for Games with Private Learning," mimeo.

Yamamoto, Y. (2007): "Efficiency Results in N Player Games with Imperfect Private Monitoring," *Journal of Economic Theory* 135, 382-413.

Yamamoto, Y. (2009): "A Limit Characterization of Belief-Free Equilibrium Payoffs in Repeated Games," *Journal of Economic Theory* 144, 802-824.

Yamamoto, Y. (2010): "Repeated Games with Private and Independent Monitoring," mimeo.