Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 12-044

"Individual Learning and Cooperation in Noisy Repeated Games"

by

Yuichi Yamamoto

http://ssrn.com/abstract=2177923

Individual Learning and Cooperation in Noisy Repeated Games*

Yuichi Yamamoto†

November 10, 2012

Abstract

We investigate whether two players in a long-run relationship can maintain cooperation when the details of the underlying game are unknown. Specifically, we consider a new class of repeated games with private monitoring, where an unobservable state of the world influences the payoff functions and/or the monitoring structure. Each player privately learns the state over time, but cannot observe what the opponent learns. We show that there are robust equilibria where players eventually obtain payoffs as if the true state were common knowledge and players played a "belief-free" equilibrium. The result is applied to various examples, including secret price-cutting with unknown demand.

Journal of Economic Literature Classification Numbers: C72, C73.

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

*This work is based on the first chapter of my Ph.D. dissertation at Harvard University. I am grateful to my advisors, Attila Ambrus, Susan Athey, and especially Drew Fudenberg for encouragement and extensive discussions. I also thank Daisuke Hirata, George Mailath, Tomasz Strzalecki, and seminar participants at various places for insightful comments.

†Department of Economics, University of Pennsylvania. Email: yyam@sas.upenn.edu

Contents

1 Introduction

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 sales are private information. In such a situation, a firm's sales level is a noisy private signal about price of the opponents, as it tends to be low if the opponents (secretly) undercut their price.[1] This is a "secret price-cutting" game of Stigler (1964), and in the literature, it is assumed that firms know the distribution of sales as a function of their price. However, in practice, firms may not know the exact distribution of sales. For example, a firm may know that there is a good chance of sales decrease if the opponents undercut their price, but may not know the exact probability of sales decrease. This is likely the case especially when firms enter a new market, as their information about the market structure is often limited. In such a case, the firms may acquire more precise information about the sales distribution through learning by doing. How do the uncertainty about the market structure and learning influence decision making by the firms? Do they have an incentive to sustain collusion in the presence of the uncertainty?

Motivated by these questions, this paper develops a general model of repeated games with *private monitoring*, where players do not know the monitoring structure. In repeated games with private monitoring, players do not directly observe their opponents' actions but instead observe noisy private signals. A secret price-cutting game is a leading example of private monitoring, and other examples include relational contracts with subjective evaluations (Levin (2003) and Fuchs (2007)) and international trade agreements in the presence of concealed trade barriers (Park (2011)). Past work has shown that a long-term relationship helps provide incentives to cooperate even under private monitoring,[2] but these results

[1] Harrington and Skrzypacz (2011) report that these properties are common to the recent lysine and vitamin markets.

[2] For example, efficiency can be approximately achieved in the prisoner's dilemma, when observations are nearly perfect (Sekiguchi (1997), Bhaskar and Obara (2002), Hörner and Olszewski (2006), Chen (2010), and Mailath and Olszewski (2011)), nearly public (Mailath and Morris (2002), Mailath and Morris (2006), and Hörner and Olszewski (2009)), statistically independent (Matsushima (2004)), or even fully noisy and correlated (Fong, Gossner, Hörner and Sannikov (2011) 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.

heavily rely on the assumption that players know the exact distribution of private signals, which is not appropriate in some economic situations, as discussed above. This paper relaxes such an assumption and examines its impact on equilibrium outcomes.

Formally, 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. 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, firms obtain higher expected payoffs at a given price at states where high sales are likely. Thus even if the payoff to each sales level is known, uncertainty about the distribution of sales yields uncertainty about the expected payoffs of the stage game.

Since observations are private in our model, players' posterior beliefs about the true state need not coincide in later periods. In particular, while each player may privately learn the state from observed signals, this learning process may not lead to "common learning" in the sense of Cripps, Ely, Mailath, and Samuelson (2008), that is, a player may *not* learn that the opponent learns the state, or a player may *not* learn that each player learns that each player learns the state, or ... For example, in the context of secret price-cutting, a firm may privately learn the true distribution of sales from its own experience, but it cannot observe the opponent's past experience and hence may not learn what the opponent learned. What happens in such a situation? Are they willing to cooperate even though they may be unsure about what the opponent learned and about what the opponent will play? The main finding of this paper is that despite the potential complications, players can still maintain some level of cooperation through appropriate use of intertemporal incentives.

In our model, to check whether a given strategy profile is sequentially rational, we need to know players' beliefs about the true state in general. However, computing these beliefs is intractable in most cases, as beliefs are updated through observed signals and there are infinitely many periods. Accordingly, characterizing the entire equilibrium set is not an easy task. Instead, we look at a tractable subset of Nash equilibria, called *belief-free ex-post equilibria* or *BFXE*. This allows us to obtain a clean characterization of the equilibrium payoff set, and as an application, we show that a large set of payoffs (including Pareto-efficient outcomes) can be achieved in many economic examples.

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 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. 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, hereafter EHO), 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.

As shown by past work, most of belief-free equilibria are mixed strategies, and players' randomization probabilities are carefully chosen to make the opponent indifferent. These indifference conditions are typically violated once the signal distribution is perturbed; as a result, the existing constructions of belief-free equilibria are not robust to a perturbation of the monitoring structure. A challenge in constructing belief-free equilibria in our setup is that we need to find randomization probabilities which satisfy all the indifference conditions even when players do not know the signal distribution and their beliefs about the signal distribution can be perturbed. If the same randomization probability satisfies the indifference conditions for all states, then it is a good candidate for an equilibrium; indeed, it constitutes a BFXE. A contribution of this paper is to identify a condition under which such a strong requirement can be satisfied and these equilibria can support a large set of non-trivial payoffs. The key is that under our informational condition, there are more possible signals than in the case of "canonical signal space" studied in the past work, which assures that there be enough room to choose appropriate randomization probabilities.

To illustrate a concrete idea of BFXE, we begin with simple examples; in Section 3.2, we consider private provision of public goods where the marginal profit from contribution is unknown and players learn it through private signals. In this situation, players cannot observe what the opponent has learned about the marginal profit; thus it is unclear how players coordinate their play in equilibrium, and as a result, various folk theorems derived in past work do not apply. We explicitly construct a BFXE and show that it attains the Pareto-efficient outcome in such an environment. Also in Section 3.3, we consider another example where players' interests are totally different at different states, and construct a BFXE. With these complete descriptions of equilibrium strategies, it is easy to see how players learn the state from private signals and use that information in BFXE. In particular, it is worth noting that the equilibrium strategies of Section 3.2 exhibit a simple form of "punish-and-forgive" behavior while those of Section 3.3 take a different simple form of "learning and adjustment" behavior, which are frequently observed in real-world activities.

Since BFXE are ex-post equilibria and players' beliefs about the state of the world are irrelevant to their best replies, one may wonder what is the value of state learning in BFXE. The key is that even though players play the same strategy profile regardless of the true state in an ex-post equilibrium, the distribution of future actions may depend on the true state because players' future play may depend on signals today, the distribution of which is influenced by the true state. In particular, there may be an ex-post equilibrium where for each state of the world, the distribution of actions conditional on that state assigns a high probability to the efficient action for that state. In this sense, state learning is valuable even if we look at ex-post equilibria.

In Section 5, we extend this idea to a general setup and obtain our main result, the state-learning theorem. It characterizes the set of BFXE payoffs with patient players under an identifiability condition, and shows that there are BFXE in which players eventually obtain payoffs as if they knew the true state and played a belief-free equilibrium for that state. This implies that BFXE can do as well as belief-free equilibria can do in the known-state game, and that the main results of EHO extend to the case where players do not know the monitoring structure. Our identifiability condition guarantees that players privately learn the true state in the long run, but does not assure that the state becomes (approximate) common knowledge; hence the result here is not an immediate consequence of the informational assumption. Applying this state-learning theorem, we show that firms can maintain collusion under a mild condition even if they do not have precise information about the market; also we show that there are BFXE approximating efficiency in many economic examples.

As argued, the set of BFXE is only a subset of Nash equilibria, and is empty for some cases (although we show that BFXE exist when players are patient and some additional conditions are satisfied; see Remark 4). 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 approach a promising direction for future research.

1.1 Literature Review

The notion of BFXE is a generalization of belief-free equilibria, which plays a central role in the study of repeated games with private monitoring. The idea of belief-free equilibria is proposed by Piccione (2002) and extended by Ely and Välimäki (2002), EHO, and Yamamoto (2007). Its limit equilibrium payoff set is fully characterized by EHO 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 (2011) proposes a generalization of belief-free equilibria, called weakly belief-free equilibria. Takahashi (2010) constructs 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.

BFXE is also related to ex-post equilibria. Some recent papers use the "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 (2011)), public monitoring and fixed states (Fudenberg and Yamamoto (2010) and Fudenberg and Yamamoto (2011a)), and changing states with an i.i.d. distribution (Miller (2012)). Note also that there are many papers working on ex-post equilibria in undiscounted repeated games; see Koren (1992) and Shalev (1994), for example.

Among these, the most closely related work is Fudenberg and Yamamoto (2010), who study the effect of uncertainty about the monitoring structure when players observe public signals rather than private signals. They look at ex-post equilibria as in this paper, and show that there are equilibria where players obtain payoffs as if they knew the state and played an equilibrium for that state. While our state-learning theorem may look similar to their result, it is not a corollary, because in our setup, public information is not available so that it is a priori unclear if ex-post equilibria and belief-free equilibria can be combined in a useful way. Indeed, in Fudenberg and Yamamoto (2010), players can form a "publicly observable dummy belief" about the true state based on public signals, which helps players coordinating their play; see Section 5.3 for more discussions. Note also that we explicitly construct equilibrium strategies in some examples and illustrate how players learn the state in ex-post equilibria. Fudenberg and Yamamoto (2010) do not have such a result.

This paper also contributes to the literature on repeated games with incomplete information. Many papers study the case where there is uncertainty about the payoff functions and actions are observable; see Forges (1984), Sorin (1984), Hart (1985), Sorin (1985), Aumann and Maschler (1995), Cripps and Thomas (2003), Gossner and Vieille (2003), Wiseman (2005), and Wiseman (2012).

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 provide 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, who might want to deviate to slow down the speed of learning. Therefore, their result does not directly apply to our setting.

2 Repeated Games with Private Learning

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, where 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 Ω. Assume that players cannot observe the true state ω, and let $\mu \in \triangle\Omega$ denote their common prior over ω.[3] Throughout the paper, we assume that the game begins with symmetric information: Each player's beliefs about ω correspond to the prior. But it is straightforward to extend our analysis to the case with asymmetric information as in Fudenberg and Yamamoto (2011a).[4]

Each period, players move simultaneously, and player $i \in I$ chooses an action a_i from a finite set A_i and observes a private signal σ_i from a finite set Σ_i.[5] Let $A \equiv \times_{i\in I} A_i$ and $\Sigma = \times_{i\in I}\Sigma_i$. The distribution of a signal profile $\sigma \in \Sigma$ depends on the state of the world ω and on an action profile $a \in A$, and is denoted by $\pi^\omega(\cdot|a) \in \triangle\Sigma$. 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 signal 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 signal 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 given an action profile

[3]Because our arguments deal only with ex-post incentives, they extend to games without a common prior. However, as Dekel, Fudenberg, and Levine (2004) argue, the combination of equilibrium analysis and a non-common prior is hard to justify.

[4]Specifically, all the results in this paper extend to the case where each player i has initial private information θ_i about the true state ω, where the set Θ_i of player i's possible private information is a partition of Ω. Given the true state $\omega \in \Omega$, player i observes $\theta_i^\omega \in \Theta_i$, where θ_i^ω denotes $\theta_i \in \Theta_i$ such that $\omega \in \theta_i$. In this setup, private information θ_i^ω allows player i to narrow down the set of possible states; for example, player i knows the state if $\Theta_i = \{(\omega_1),\cdots,(\omega_o)\}$. For games with asymmetric information, we can allow different types of the same player to have different best replies as in PTXE of Fudenberg and Yamamoto (2011a); to analyze such equilibria, regime R should specify recommended actions for each player i and each type θ_i, i.e., $R = R = (R_i^{\theta_i})_{(i,\theta_i)}$.

[5]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).

a.[6]

As emphasized in the introduction, uncertainty about the payoff functions and/or the monitoring structure is common in applications. Examples that fit our model include secret price-cutting with unknown demand function and moral hazard with subjective evaluation and unknown evaluation distribution. Also a repeated game with observed actions and individual learning is a special case of the above model. To see this, let $\Sigma_i = A \times Z_i$ for some finite set Z_i and 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 $a'' \neq a$. Under this setup, actions are perfectly observable by players (as σ_i must be consistent with the action profile a) and players learn the true state ω from private signals z_i. More concrete examples will be given in the next section.

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 private histories h_i^t. Also, we 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 history profiles h^t. A strategy for player i is defined to be a mapping $s_i : \bigcup_{t=0}^\infty H_i^t \to \triangle A_i$. 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

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

that is, $V(\omega)$ is the set of the convex hull of possible stage-game payoff vectors given ω. Then 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., $v = ((v_1^\omega, v_2^\omega))_{\omega \in \Omega}$. 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

[6]If there are $\omega \in \Omega$ and $\tilde{\omega} \neq \omega$ such that $u_i^\omega(a_i, \sigma_i) \neq u_i^{\tilde{\omega}}(a_i, \sigma_i)$ 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.

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.

3 Motivating Examples

In this section, we consider a series of examples to illustrate the scope of our model and the idea of our equilibrium strategies when players learn the true state from private signals.

3.1 Secret Price Cutting

Suppose that there are two firms in a market. The firms do not know the true state $\omega \in \Omega$ and they have a common prior $\mu \in \triangle\Omega$. In every period, firm i chooses its price $a_i \in A_i$. Firm i's sales level $y_i \in Y_i$ depends on the price vector $a = (a_1, a_2)$ and an unobservable aggregate shock $\eta \in [0,1]$, which follows a distribution $F^{\omega}(\cdot|a)$ with density $f^{\omega}(\cdot|a)$. Given (a,η), we denote the corresponding sales level of firm i by $y_i(a,\eta)$. Firm i's profit is $u_i(a_i, y_i) = a_i y_i - c_i(y_i)$ where $c_i(y_i)$ is the production cost. In this setup, the distribution of sales level profile $y = (y_1, y_2)$ conditional on (ω, a) is given by $\pi^{\omega}(\cdot|a)$, where $\pi^{\omega}(y|a) = \int_{\eta\in\{\tilde{\eta}|y=(y_1(a,\tilde{\eta}),y_2(a,\tilde{\eta}))\}} f^{\omega}(\eta|a)d\eta$. Also, firm i's expected payoff at state ω given a is $g_i^{\omega}(a) = \sum_{y_i\in Y_i} \pi^{\omega}(y|a)u_i(a_i, y_i)$.

Rotemberg and Saloner (1986) consider a repeated duopoly model where an aggregate shock η is observable to the firms and follows an i.i.d. process. The model here differs from theirs in that (i) an aggregate shock is not observable and (ii) its distribution is unknown to the firms. This is a natural assumption in some economic situations; for example, when the firms enter a new market, they may not know the structure of the market and hence may not know the exact distribution of an aggregate shock. This is one of the leading examples of our general model introduced in Section 2. In Section 5.1 we will apply our main result to this example and give a condition under which the firms can sustain collusion even if they do not know the distribution of sales.

In this example, the utility function and signal distribution have very general forms, and accordingly it is hard to illustrate the idea of our equilibrium construction. In the next two subsections, we consider simpler examples where actions are perfectly observable, and describe how to construct equilibrium strategies when players learn the state from private signals. Here we stress that we assume observable actions just to make our exposition as simple as possible. Indeed, as will be explained, a similar construction is valid even if players observe noisy information about actions.

3.2 Private Provision of Public Goods

There are two players and two possible states, so $\Omega = \{\omega_1, \omega_2\}$. In each period t, each player i makes a decision on whether to contribute to a public good or not. Let $A_i = \{C_i, D_i\}$ be the set of player i's possible actions, where C_i means contributing to a public good and D_i means no contribution. After making a decision, each player i receives a stochastic output z_i from a finite set Z_i. An output z_i is private information of player i and its distribution depends on the true state ω and on the total investment $a \in A$. Note that many economic examples fit this assumption, as firms' profits are often private information and firms are often uncertain about the distribution of profits. We also assume that a choice of contribution levels is perfectly observable to players; 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 $a'' \neq a$. With an abuse of notation, let $\pi^\omega(z|a)$ denote the joint distribution of $z = (z_1, z_2)$ given (a, ω); that is, $\pi^\omega(z|a) = \pi^\omega((a, z_1), (a, z_2)|a)$. We do not impose any assumption on the joint distribution of (z_1, z_2), so that outputs z_1 and z_2 can be independent or correlated. When z_1 and z_2 are perfectly correlated, our setup reduces to the case where outputs are public information.

Player i's actual payoff does not depend on the state ω and is given by $u_i(a_i, \sigma_i) = \tilde{u}_i(z_i) - c_i(a_i)$, where $\tilde{u}_i(z_i)$ is player i's profit from an output z_i and $c_i(a_i)$ is cost of contributions. We assume $c_i(C) > c_i(D) = 0$, that is, contribution is costly.

As in the general model introduced in Section 2, the expected payoff of firm i at state ω is denoted by $g_i^\omega(a) = \sum_{\sigma \in \Sigma} \pi^\omega(\sigma|a) u_i(a_i, \sigma_i)$. Note that a player's expected payoff depends on the true state ω, as it influences the distribution of outputs z. We assume that the expected payoffs are as in the following tables:

	C	D
C	3, 3	−1, 4
D	4, −1	0, 0

	C	D
C	3, 3	1, 4
D	4, 1	0, 0

The left table denotes the expected payoffs for state ω_1, and the right table for state ω_2. Note that the stage game is a prisoner's dilemma at state ω_1, and is a chicken game at state ω_2. This captures the situation where contributions are socially efficient but players have a free-riding incentive; indeed, in each state, (C,C) is efficient but a player is willing to choose D when the opponent chooses C. Another key feature of this payoff function is that players do not know the marginal benefit from contributing to a public good and do not know whether they should contribute, given that the opponent does not contribute. Specifically, the marginal benefit is low in ω_1 so that a player prefers D to C when the opponent chooses D, while the marginal profit is high in ω_2 so that a player prefers C.

Since actions are observable, one may expect that the efficient payoff vector $((3,3),(3,3))$ can be approximated by standard trigger strategies. But this approach does not work, because there is no static ex-post equilibrium in this game and how to punish a deviator is not obvious. Note also that the folk theorems of Fudenberg and Yamamoto (2010) and Wiseman (2012) do not apply here, as they assume that players obtain public (or almost public) information about the true state in each period. In this example, players learn the true state ω only through private information z_i and it is unclear whether players are willing to cooperate after learning the true state ω.

In what follows, we will construct a simple equilibrium strategy with payoff $((3,3),(3,3))$, assuming that players are patient. We assume that for each i and $a \in A$, there are outputs $z_i^{\omega_1}(a)$ and $z_i^{\omega_2}(a)$ such that

$$\frac{\pi_i^{\omega_1}(z_i^{\omega_1}(a)|a)}{\pi_i^{\omega_2}(z_i^{\omega_1}(a)|a)} \geq 2 \quad \text{and} \quad \frac{\pi_i^{\omega_2}(z_i^{\omega_2}(a)|a)}{\pi_i^{\omega_1}(z_i^{\omega_2}(a)|a)} \geq 2 \tag{1}$$

where $\pi_i^{\omega}(\cdot|a)$ is the marginal distribution of z_i given (ω,a). That is, the marginal distributions of z_i are sufficiently different at different states, so that given any action profile a, there is an output level z_i^{ω} that has a sufficiently high likelihood ratio to test for the true state being ω. This assumption is not necessary for the existence of asymptotically efficient equilibria (see Section 5.2 for details), but it considerably simplifies our equilibrium construction, as shown below.

In our equilibrium, each player uses a strategy which is implemented by a two-state automaton. Specifically, player 1 uses the following strategy:

States: Given any period t and given any history h_1^t, player 1 is in one of the two states, either $x(1)$ or $x(2)$. In state $x(1)$, player 1 chooses C to "reward" player 2. In state $x(2)$, player 1 chooses D to "punish" player 2.

Transition after State *x*(1): Suppose that player 1 is currently in the reward state $x(1)$ so that she chooses C today. For the next period, player 1 will switch to the punishment state $x(2)$ with some probability depending on today's outcome. Specifically, given player 2's action $a_2 \in A_2$ and player 1's output $z_1 \in Z_1$, player 1 will go to the punishment state $x(2)$ with probability $\beta(a_2, z_1)$ and stay at the reward state $x(1)$ with probability $1 - \beta(a_2, z_1)$. We set $\beta(C, z_1) = 0$ for all z_1; that is, player 1 will reward player 2 for sure if player 2 chooses C today. $\beta(D, z_1)$ will be specified later, but we will have $\beta(D, z_1) > 0$ for all z_1, that is, player 1 will punish player 2 with positive probability if player 2 chooses D today.

Transition after State *x*(2): Suppose that player 1 is in the punishment state $x(2)$ so that she chooses D today. For the next period, player 1 will switch to the reward state $x(1)$ with some probability depending on today's outcome. Specifically, given (a_2, z_1), player 1 will go to $x(1)$ with probability $\gamma(a_2, z_1)$ and stay at $x(2)$ with probability $1 - \gamma(a_2, z_1)$. $\gamma(a_2, z_1)$ will be specified later, but we will have $\gamma(a_2, z_1) > 0$ for all a_2 and z_1, that is, player 1 will switch to the reward state with positive probability no mater what player 2 does.

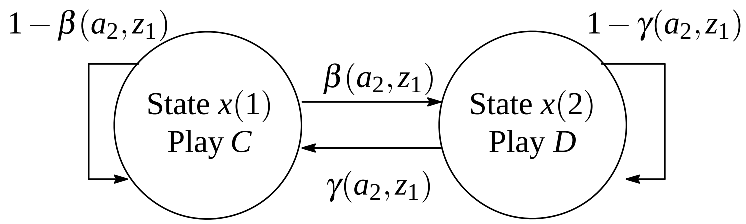

Figure 1: Automaton

The equilibrium strategy here is simple and intuitive. Consider the automaton with the initial state $x(1)$. This strategy asks a player to cooperate until the opponent deviates, and to switch to the punishment phase $x(2)$ after a deviation. In the punishment phase, she chooses D to punish the opponent, and then returns to the cooperative phase $x(1)$ to forgive the opponent. This "punish-and-forgive" behavior is commonly observed in the real world.

In what follows, we will show that this punish-and-forgive behavior actually constitutes an equilibrium if we choose the transition probabilities carefully. The key idea here is to choose player 1's transition probabilities β and γ in such a way that player 2 is indifferent between C and D regardless of player 1's current state of the automaton and of the state of the world ω. This means that player 2 is indifferent between C and D after every history, so that any strategy is a best response to player 1's strategy. Also, we construct player 2's strategy in the same way so that player 1 is always indifferent between C and D. Then a pair of such strategies constitutes an equilibrium, as they are best replies to each other. An advantage of this equilibrium construction is that a player's best reply is independent of her belief about the state of the world ω and of her belief about the opponent's history, so that we do not need to compute these beliefs to check its incentive compatibility. We call such a strategy profile *belief-free ex-post equilibrium* (*BFXE*).

More specifically, we will choose the transition probabilities in such a way that the following properties are satisfied:

- If player 1 is currently in the reward state $x(1)$, then player 2's continuation payoff from today is 3 given any state of the world ω, no matter what player 2 plays.

- If player 1 is currently in the punishment state $x(2)$, then player 2's continuation payoff is 2 at ω_1 and $\frac{7}{3}$ at ω_2, no matter what player 2 plays.

For each $k = 1,2$, let $v_2(k)$ denote the target payoff vector of player 2 given player 1's state $x(k)$; that is, $v_2(1) = (v_2^{\omega_1}(1), v_2^{\omega_2}(1)) = (3,3)$ and $v_2(2) = (v_2^{\omega_1}(2), v_2^{\omega_2}(2)) = (2, \frac{7}{3})$, where $v_2^{\omega}(k)$ is the target payoff given $x(k)$ and ω. Figure 2 describes these target payoffs and stage game payoffs. The horizontal axis denotes player 2's payoff at ω_1, and the vertical axis denotes player 2's payoff at ω_2. The point $(4,4)$ is the payoff vector of the stage game when (C,D) is played. Likewise, the points

$(3,3)$, $(-1,1)$, and $(0,0)$ are generated by (C,C), (D,C), and (D,D), respectively. The bold line is the convex hull of the set of target payoff vectors, $v_2(1)$ and $v_2(2)$.

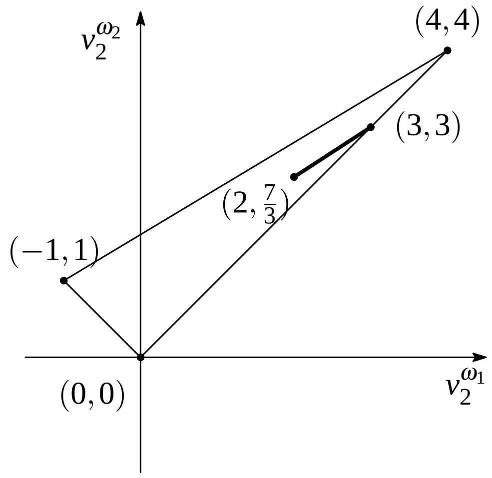

Figure 2: Payoffs

When the discount factor δ is close to one, there indeed exist the transition probabilities β and γ such that these target payoffs are exactly achieved. To see this, consider the case where player 1 is currently in the reward state $x(1)$ so that the target payoff is $v_2(1) = (3,3)$. If player 2 chooses C, then player 2's stage-game payoff today is 3 regardless of ω, which is exactly equal to the target payoff; hence player 1 will stay at the state $x(1)$ for the next period, i.e., we set $\beta(C,z_1) = 0$ for all z_1. On the other hand, if player 2 chooses D, then player 2's stage-game payoff is 4 regardless of ω, which is higher than the target payoff. To offset this difference, player 1 will switch to the punishment state $x(2)$ with positive probability $\beta(D,z_1) > 0$ for the next period. Here the transition probability $\beta(D,z_1) > 0$ is carefully chosen so that the instantaneous gain by playing D and the expected loss by the future punishment cancel out; i.e., we choose $\beta(D,z_1)$ such that

$$g_2^{\omega}(C,D) - v_2^{\omega}(1) = \frac{\delta}{1-\delta}\sum_z \pi^{\omega}(z|C,D)\beta(D,z_1)(v_2^{\omega}(1) - v_2^{\omega}(2)) \qquad (2)$$

for each ω. (Note that the left-hand side is the instantaneous gain by playing D while the right-hand side is the expected loss in continuation payoffs.) The formal proof of the existence of such β is given in Appendix A, but the basic idea is to let $\beta(D, z_1^{\omega_2}(C,D)) > \beta(D, z_1)$ for all $z_1 \neq z_1^{\omega_2}(C,D)$, that is, we ask player 1 to switch to the punishment state with a high probability when the output level is $z_1^{\omega_2}(C,D)$, which is an indication of ω_2. The intuition is that the punishment by switching to the state $x(2)$ is less harsh at ω_2 than at ω_1 (note that $v_2^{\omega_2}(2) > v_2^{\omega_1}(2)$), and hence player 1 needs to switch to $x(2)$ more likely at ω_2 to offset player 2's instantaneous gain.

Likewise, consider the case where player 1 is in the punishment state $x(2)$ so that the target payoff is $v_2(2) = (2, \frac{7}{3})$. In this case, player 1 chooses D, so that no matter what player 2 does, player 2's stage-game payoff is lower than the target payoff regardless of ω. So player 1 will switch to the reward state $x(1)$ with positive probability $\gamma(a_2, z_1) > 0$ to offset this difference. The proof of the existence of such γ is very similar to that of β and is found in Appendix A.

With such a choice of β and γ, player 2 is always indifferent between C and D and hence any strategy of player 2 is a best reply. Also, as explained, we construct player 2's two-state automaton in the same way, so that player 1 is always indifferent. Then the pair of these strategies constitutes an equilibrium. In particular, when both players begin their play from the reward state $x(1)$, its equilibrium payoff is $((3,3),(3,3))$, as desired.

In this efficient equilibrium, players choose (C,C) forever unless somebody deviates to D. If player $-i$ deviates and chooses D, player i punishes this deviation by switching to $x(2)$ with positive probability and starting to play D. However, "always play D" is too harsh compared to the target payoff $(2, \frac{7}{3})$, and hence player i comes back to $x(1)$ with some probability after every period. In the long run, both players come back to $x(1)$ and play (C,C), because this is the unique absorbing state of the automaton.

The above two-state automaton is a generalization of that of Ely and Välimäki (2002) for a repeated prisoner's dilemma with almost-perfect monitoring. The reason why their equilibrium construction directly extends is that in this example, the payoffs at different states are "similar" in the sense that for each ω, the action C can be used to reward the opponent and D to punish. When this structure is lost, a player is not sure about what action should be taken to reward or punish the opponent, so that state learning becomes more important. In the next example, we show how the equilibrium strategies look like in such environments.

3.3 Conflicting Interests

There are two players and two possible states, ω_1 and ω_2. In each stage game, player 1's action is either U or D, and player 2's action is either L or R. They choose actions simultaneously in each stage game.

In state ω_1, the stage game is a prisoner's dilemma; (U,L) is efficient but D strictly dominates U and R strictly dominates L. In state ω_2, the stage game is also a prisoner's dilemma, but the role of the actions are reversed. That is, (D,R) is efficient and (U,L) is a Nash equilibrium. The following tables summarize this payoff structure:

	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 efficient payoff vector $((1,1),(1,1))$ is not feasible in a one-shot game, as players need to choose different action profiles at different states to generate this payoff. (They need to play (U,L) at ω_1 and (D,R) at ω_2.)

Suppose that each player i observes an action profile and a noisy private signal $z_i \in Z_i = \{z_i^{\omega_1}, z_i^{\omega_2}\}$ about the true state in every period. The probability of $z_i^{\omega_1}$ is $\frac{2}{3}$ at state ω_1 and $\frac{1}{3}$ at state ω_2, regardless of which actions players play. Likewise, the probability of $z_i^{\omega_2}$ is $\frac{2}{3}$ at state ω_2 and $\frac{1}{3}$ at state ω_1, regardless of actions. So the signal $z_i^{\omega_1}$ indicates that the true state is likely to be ω_1 and $z_i^{\omega_2}$ means that the true state is likely to be ω_2. Again, this likelihood ratio assumption is not necessary for the existence of asymptotically efficient equilibria , but it simplifies our equilibrium construction. We impose no assumption on the joint distribution of z_1 and z_2, so these signals can be independent or correlated.

In this example, the payoff functions are totally different at different states, so that state learning is necessary to provide proper intertemporal incentives. However, since players learn the true state from private signals, they may not know what the opponent has learned in the past play and it is unclear how players create such incentives. Our goal is to give a simple and explicit equilibrium construction where players learn the state and adjust their actions. As in the previous example, our equilibrium is a BFXE, that is, each player is indifferent between the two actions given any history and given any state of the world ω.

We will construct an equilibrium where player 1 (player 2) tries to learn the true state ω from private signals at the beginning, and then adjust her play to choose an "appropriate" action; she chooses U (L) when she believes that the true state is ω_1, and she chooses D (R) when she believes that the true state is ω_2. Specifically, player 1's strategy is described by the following four-state automaton:

States: Given any period t and after any history, player 1 is in one of the four states, $x(1)$, $x(2)$, $x(3)$, or $x(4)$. Player 1 chooses U in states $x(1)$ and $x(2)$, while she chooses D in states $x(3)$ and $x(4)$. As in the previous example, we denote by $v_2(k) = (v_2^{\omega_1}(k), v_2^{\omega_2}(k))$ player 2's ex-post payoffs of the repeated game when player 1's play begins with the state $x(k)$. Set

$$\begin{aligned}
v_2(1) &= (v_2^{\omega_1}(1), v_2^{\omega_2}(1)) = (1,0),\\
v_2(2) &= (v_2^{\omega_1}(2), v_2^{\omega_2}(2)) = (0.8, 0.79),\\
v_2(3) &= (v_2^{\omega_1}(3), v_2^{\omega_2}(3)) = (0.79, 0.8),\\
v_2(4) &= (v_2^{\omega_1}(4), v_2^{\omega_2}(4)) = (0,1).
\end{aligned}$$

Roughly, player 1 is in state $x(1)$ when she believes that the true state is ω_1 and wants to reward player 2; thus player 2's target payoff is high at ω_1 ($v_2^{\omega_1}(1) = 1$), but it is low at ω_2 ($v_2^{\omega_1}(1) = 0$). Likewise, player 1 is in state $x(4)$ when she believes that the true state is ω_2 and wants to reward player 2. In states $x(2)$ and $x(3)$, player 1 is still unsure about ω; she moves back and forth between these two states for a while, and after learning the true state ω, she moves to $x(1)$ or $x(4)$. The detail of the transition rule is specified below, but intuitively, when player 1 gets convinced that the true state is ω_1, she will move to $x(1)$ and choose the appropriate action U. Likewise, when player 1 becomes sure that the true state is ω_2, she will move to $x(4)$ and choose D. This "learning and adjustment" allows player 2 to obtain high expected payoffs at both ω_1 and ω_2 when player 1 is begins her play from $x(2)$ or $x(3)$, as shown in Figure 3.

Transitions after *x*(1): If player 2 chooses L today, then player 1 stays at $x(1)$ for sure. If player 2 chooses R today, then player 1 switches to $x(4)$ with probability $\frac{1-\delta}{\delta}$, and stays at $x(1)$ with probability $1 - \frac{1-\delta}{\delta}$.

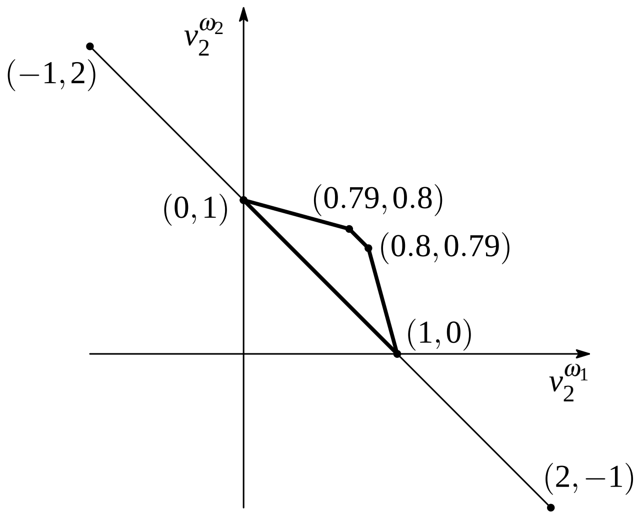

Figure 3: Payoffs

The idea of this transition rule is the following. When player 2 chooses L, the stage game payoff for player 2 is $(1,0)$, which is exactly the target payoff $v_2(1)$, so that player 1 stays at $x(1)$ for sure. On the other hand, when player 2 chooses R, the stage game payoff for player 2 is $(2,-1)$, which is different from the target payoff. In this case player 1 moves to $x(4)$ with positive probability to offset this difference.

Transitions after $x(2)$: Suppose that player 2 chooses L today. If $z_1^{\omega_1}$ is observed, player 1 goes to $x(1)$ with probability $\frac{(1-\delta)117}{\delta}$ and stays at $x(2)$ with the remaining probability. If $z_1^{\omega_2}$ is observed, then she goes to $x(3)$ with probability $\frac{(1-\delta)4740}{\delta}$ and stays at $x(2)$ with the remaining probability. That is, player 1 moves to $x(1)$ only when she observes $z_1^{\omega_1}$ and gets more convinced that the true state is ω_1.

Suppose next that player 2 chooses R today. In this case, if $z_1^{\omega_1}$ is observed, player 1 goes to $x(3)$ with probability $\frac{(1-\delta)61}{\delta}$ and stays at $x(2)$ with the remaining probability. If $z_1^{\omega_2}$ is observed, then she goes to $x(3)$ with probability $\frac{(1-\delta)238}{\delta}$ and stays at $x(2)$ with the remaining probability. Note that player 1 will not move to $x(1)$ in this case regardless of her signal z_1. The reason is that when player 2 chooses R, her stage-game payoff at ω_1 is 2, which is too high compared to the target payoff; to offset this difference, player 1 needs to give lower continuation payoffs to player 2 by moving to $x(3)$ rather than $x(1)$.

Transitions after x(3): The transition rule is symmetric to the one after $x(2)$. Suppose that player 2 chooses R today. If $z_1^{\omega_2}$ is observed, player 1 goes to $x(4)$ with probability $\frac{(1-\delta)117}{\delta}$ and stays at $x(3)$ with the remaining probability. If $z_1^{\omega_1}$ is observed, then she goes to $x(2)$ with probability $\frac{(1-\delta)4740}{\delta}$ and stays at $x(3)$ with the remaining probability.

Suppose next that player 2 chooses L today. If $z_1^{\omega_2}$ is observed, player 1 goes to $x(2)$ with probability $\frac{(1-\delta)61}{\delta}$ and stays at $x(3)$ with the remaining probability. If $z_1^{\omega_1}$ is observed, then she goes to $x(2)$ with probability $\frac{(1-\delta)238}{\delta}$ and stays at $x(3)$ with the remaining probability.

Transitions after x(4): The transition rule is symmetric to the one after $x(1)$. If player 2 chooses R today, then stay at $x(4)$. If player 2 chooses L today, then go to $x(1)$ with probability $\frac{1-\delta}{\delta}$, and stay at $x(4)$ with probability $\frac{2\delta-1}{\delta}$

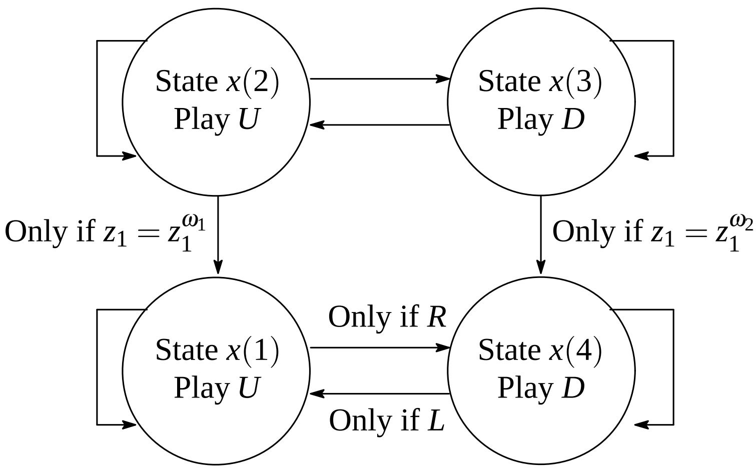

Figure 4: Automaton

Simple algebra (like (2) in the previous example) shows that given any ω and $x(k)$, player 2 is indifferent between L and R and her overall payoff is exactly $v_2^{\omega}(k)$. This means that any strategy of the repeated game is optimal for player 2, and in particular, if player 1's initial state is $x(2)$, then player 2's overall payoff is $(0.8, 0.79)$. We can construct player 2's automaton in the same way, and it is easy to see that a pair of these automata is an equilibrium of the repeated game. When both players begin with initial state $x(2)$, the equilibrium payoff is $((0.8, 0.8), (0.79, 0.79))$, which cannot be achieved in a one-shot game. This example shows that BFXE is useful even when the payoff functions are totally different at different states. Note that the equilibrium strategy here takes a simple form of "learning and adjustment," which is seemingly a reasonable play when there is uncertainty of the underlying payoff structure. What we found is that if the adjustment process is carefully chosen, it actually constitutes an equilibrium.

A natural question here is if there are more efficient equilibria, and in particular, if we can approximate the payoff vector $((1,1),(1,1))$. The main reason why our equilibrium payoff is bounded away from $((1,1),(1,1))$ is that although players can obtain precise information about the true state ω in the long run through private signals, they do not use that information in an efficient way. Too see this, note that in the above equilibrium, a player's continuation strategy depends only on the current state of the automaton and today's outcome; that is, private signals in the past play can influence a player's continuation play only through the current state of the automaton. But there are only four possible states ($x(1)$, $x(2)$, $x(3)$, or $x(4)$) in the automaton, which means that they are less informative about ω than the original private signals. (In other words, the state of the automaton can represent only coarse information about ω.) Accordingly, players fail to play an efficient action with a non-negligible probability. For example, even if the true state is ω_1, the probability that she reaches the state $x(4)$ in the long run conditional on the true state ω_1 and the initial state $x(2)$ is bounded away from zero.

This problem can be solved by considering an automaton with more states; if we increase the number of states of the automaton, then information classification becomes finer, which allows us to construct more efficient equilibria. For example, there is an automaton with six states which generates the following payoffs:[7]

[7] These payoffs are generated by the following automaton:

Actions: Player 1 chooses U in states $x(1)$, $x(2)$, and $x(3)$ and D in states $x(4)$, $x(5)$, and $x(6)$.

Transitions after $x(1)$: If player 2 chooses L today, then player 1 stays at $x(1)$ for sure. If player 2 chooses R today, then player 1 switches to $x(6)$ with probability $\frac{1-\delta}{\delta}$, and stays at $x(1)$ with probability $1 - \frac{1-\delta}{\delta}$.

Transitions after $x(2)$: Suppose that player 2 chooses L today. If $z_1^{\omega_1}$ is observed, player 1 goes to $x(1)$ with probability $\frac{(1-\delta)39}{\delta}$ and stays at $x(2)$ with the remaining probability. If $z_1^{\omega_2}$ is observed, then she goes to $x(3)$ with probability $\frac{(1-\delta)1890}{\delta}$ and stays at $x(2)$ with the remaining

$$v_2(1) = (v_2^{\omega_1}(1), v_2^{\omega_2}(1)) = (1, 0),$$
$$v_2(2) = (v_2^{\omega_1}(2), v_2^{\omega_2}(2)) = (0.93, 0.9),$$
$$v_2(3) = (v_2^{\omega_1}(3), v_2^{\omega_2}(3)) = (0.927, 0.91),$$
$$v_2(4) = (v_2^{\omega_1}(4), v_2^{\omega_2}(4)) = (0.91, 0.927),$$
$$v_2(5) = (v_2^{\omega_1}(5), v_2^{\omega_2}(5)) = (0.9, 0.93),$$
$$v_2(6) = (v_2^{\omega_1}(6), v_2^{\omega_2}(6)) = (0, 1).$$

We can show that as we increase the number of states of an automaton, more efficient payoffs are achievable and the efficient payoff $((1,1),(1,1))$ are eventually approximated.[8] Also there are asymptotically efficient equilibria even when we consider a general signal distribution; see Section 5.2 for more details.

Remark 1. In this section, we have looked at games with observed actions, but a similar equilibrium construction applies to games with private and almost-perfect monitoring, where each player does not observe actions directly but receives private information about actions with small noise. The idea is that even if small noise is introduced to the monitoring structure, we can slightly perturb the target payoffs $\{v_i(k)\}$ and the transition probabilities so that the resulting automaton is still an equilibrium. The logic is very similar to the one for belief-free equilibria (Ely and Välimäki (2002) and EHO) and hence omitted.

---

probability. Suppose next that player 2 chooses R today. In this case, if $z_1^{\omega_1}$ is observed, player 1 goes to $x(3)$ with probability $\frac{(1-\delta)1570}{\delta 3}$ and stays at $x(2)$ with the remaining probability. If $z_1^{\omega_2}$ is observed, then she goes to $x(3)$ with probability $\frac{(1-\delta)70}{\delta 3}$ and stays at $x(2)$ with the remaining probability.

Transitions after $x(3)$: Suppose that player 2 chooses L today. If $z_1^{\omega_1}$ is observed, player 1 goes to $x(2)$ with probability $\frac{(1-\delta)1146}{\delta}$ and stays at $x(3)$ with the remaining probability. If $z_1^{\omega_2}$ is observed, then she goes to $x(4)$ with probability $\frac{(1-\delta)7095}{\delta 17}$ and stays at $x(3)$ with the remaining probability. Suppose next that player 2 chooses R today. In this case, if $z_1^{\omega_1}$ is observed, player 1 goes to $x(4)$ with probability $\frac{(1-\delta)236}{\delta 17}$ and stays at $x(3)$ with the remaining probability. If $z_1^{\omega_2}$ is observed, then she goes to $x(4)$ with probability $\frac{(1-\delta)2747}{\delta 17}$ and stays at $x(3)$ with the remaining probability.

The specification of the transitions after $x(4)$, $x(5)$, $x(6)$ is symmetric so that we omit it.

[8] The formal proof is available upon request.

4 Belief-Free Ex-Post Equilibrium

In the previous section, we have constructed equilibrium strategies where each player is indifferent over all actions given any state of the world ω and given any past history of the opponent. An advantage of this equilibrium construction is that we do not need to compute a player's belief for checking the incentive compatibility, which greatly simplifies the analysis.

In this section, we generalize this idea and introduce a notion of *belief-free ex-post equilibria*, which is a special class of Nash equilibria. Given a strategy $s_i \in S_i$, let $s_i|_{h_i^t}$ denote the continuation strategy induced by s_i when player i's past private history was $h_i^t \in H_i^t$. For notational convenience, let $s|_{h^t}$ denote the continuation strategy profile given a history h^t, i.e., $s|_{h^t} = (s_i|_{h_i^t})_{i \in I}$.

Definition 1. A strategy profile $s \in S$ is a *belief-free ex-post equilibrium* or *BFXE* if $s_i|_{h_i^t}$ is a best reply to $s_{-i}|_{h_{-i}^t}$ in the infinitely repeated game with the true state ω for each $i \in I$, $\omega \in \Omega$, $t \geq 0$, and $h^t \in H^t$.

In BFXE, a player's best reply does not depend on the true state or the opponent's private history, so that her belief about the state and the past history is payoff-irrelevant. Thus we do not need to compute these beliefs for the verification of incentive compatibility, which exactly captures the main idea of the equilibrium construction in the previous section. BFXE reduces to belief-free equilibria of EHO in known-state games where $|\Omega| = 1$. Note that repetition of a static ex-post equilibrium is a BFXE. Note also that BFXE may not exist; for example, if there is no static ex-post equilibrium and the discount factor is close to zero, then there is no BFXE.

Given a BFXE s, let $R_i^t \subseteq A_i$ denote the set of all (ex-post) optimal actions for player i in period t, i.e., R_i^t is the set of all $a_i \in A_i$ such that $\tilde{s}_i(h_i^0) = a_i$ for some $\tilde{s}_i \in S_i$ such that $\tilde{s}_i$ is a best reply to $s_{-i}|_{h_{-i}^{t-1}}$ given any h_{-i}^{t-1} and ω. Let $R^t = \times_{i \in I} R_i^t$, and we call the set R^t the *regime for period* t. Note that the regime R^t is non-empty for any period t; indeed, if an action a_i is played with positive probability after some history h_i^{t-1}, then by the definition of BFXE, a_i is an element of R_i^t. The equilibrium strategies in the previous section are a special class of BFXE such that the corresponding regimes are given by $R^t = A$ for all t, that is, all actions are ex-post optimal in every period. For other BFXE, R_i^t can be a strict subset of A_i;

e.g., when the stage game has a strict ex-post equilibrium a, playing a in every period is a BFXE of the repeated game, and it induces the regime sequence such that $R_i^t = \{a_i\}$ for all i and t. Let $\mathscr{R}$ be the set of all possible regimes, i.e.,

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

EHO show that allowing access to public randomization 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 $y \in Y$ at the beginning of every period, where Y is the set of possible public signals. Public signals are i.i.d. draws from the same distribution $p \in \triangle Y$. 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 $s_i : \bigcup_{t=0}^\infty (H_i^t \times Y) \to \triangle A_i$. Let $s_i|_{(h_i^t, y^{t+1})}$ denote the continuation strategy of player i when her history up to period t was h_i^t and the public signal at the beginning of period $t+1$ was y^{t+1}.

As in EHO, we consider the case where $Y = \mathscr{R}$; this is the case where a public signal y suggests a regime in each period. Let S_i^* denote the set of all strategies s_i such that player i chooses her action from a suggested regime in each period. That is, S_i^* is the set of all s_i such that $\sum_{a_i\in R_i} s_i(h_i^{t-1}, R)[a_i] = 1$ for each t, h_i^{t-1}, and R.

Definition 2. Given a public randomization $p \in \triangle\mathscr{R}$, a strategy profile $s \in S$ is a *stationary BFXE with respect to p* (or *BFXE with respect to p* in short) if (i) $s_i \in S_i^*$ for each i and (ii) $\tilde{s}_i|_{h_i^0, R}$ is a best reply to $s_{-i}|_{(h_{-i}^{t-1}, R)}$ in the infinitely repeated game with the true state ω for each i, ω, t, h_{-i}^{t-1}, R and $\tilde{s}_i \in S_i^*$.

Clause (i) says that each player i chooses her action from a suggested regime in each period. Clause (ii) says that choosing a recommended action is optimal given any state ω and given any past history (h_{-i}^{t-1}, R). In a stationary BFXE, a regime is randomly chosen according to the same distribution in each period; this recursive structure allows us to use dynamic programming techniques for the analysis.

An important feature of stationary BFXE is that in the limit as $\delta \to 1$, the payoff set of BFXE without public randomization is equal to the union of the sets of stationary BFXE payoffs over all $p \in \triangle\mathscr{R}$. This is true because public randomization can substitute any regime sequence $\{R^t\}_{t=1}^{\infty}$ induced by BFXE without public randomization.[9] This feature means that characterizing the limit payoff set of BFXE without public randomization reduces to computing the limit set of stationary BFXE payoffs for each p. In what follows, we will characterize the set of stationary BFXE payoffs.

Given a discount factor $\delta \in (0,1)$ and given $p \in \triangle\mathscr{R}$, let $E^p(\delta)$ denote the set of BFXE payoffs with respect to p, 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 a stationary BFXE s with respect to p satisfying $(1-\delta)E[\sum_{t=1}\delta^{t-1}g_i^{\omega}(a^t)|s,\omega,p] = v_i^{\omega}$ for all i and ω. Note that $v \in E^p(\delta)$ specifies the equilibrium payoff for all players and for all possible states. Also, for each i, let $E_i^p(\delta)$ denote the set of player i's BFXE payoffs with respect to p, i.e., $E_i^p(\delta)$ is the set of all $v_i = (v_i^{\omega})_{\omega\in\Omega}$ 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 ω.

The following proposition asserts that given public randomization p, stationary BFXE 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 $\tilde{s}_{-i}|_{\tilde{h}^t_{-i}}$ for any t and $\tilde{h}^t_{-i}$, and thus playing $s_i|_{h_i^t}$ is a best reply to $\tilde{s}_{-i}|_{\tilde{h}^t_{-i}}$ for any t, h_i^t, and $\tilde{h}^t_{-i}$. Likewise $\tilde{s}_i|_{h_i^t}$ is a best reply to $s_{-i}|_{\tilde{h}^t_{-i}}$ for any t, h_i^t, and $\tilde{h}^t_{-i}$. 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 public randomization p, the equilibrium payoff set has a product structure. This conclusion follows from the fact that stationary BFXE are interchangeable.

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

[9]The proof is very similar to the on-line appendix of EHO and hence omitted.

Proof. To see this, fix $p \in \triangle\mathscr{R}$, 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. *Q.E.D.*

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. This idea is formalized as "individual ex-post self-generation" in Appendix D.1.1, which is useful to establish Proposition 3 in the next section.

Remark 2. One implication from the interchangeability of BFXE is that each player is willing to play an equilibrium strategy even if she does not have the correct belief about the opponent's strategy. To be precise, given $p \in \triangle\mathscr{R}$, let $S^p = \times_{i \in I} S_i^p$ be the set of BFXE strategy profiles, where S_i^p be the set of BFXE strategies of player i. Then Proposition 1 says that player i is willing to play an equilibrium strategy $s_i \in S_i^p$ as long as the opponent chooses her strategy from the set S_{-i}^p. That is, players are willing to play a BFXE as long as they have correct beliefs about the randomization $p \in \triangle\mathscr{R}$ of regimes, but not about strategies.

Remark 3. 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).

5 State-Learning Theorem

5.1 General Case

In Section 3, we have focused on some examples and shown that there are BFXE where players learn the true state from private signals and adjust their continuation play. In this section, we extend the analysis to a general setup, and show that if a certain informational condition is satisfied, the set of BFXE payoffs in the limit as $\delta \to 1$ is equal to the product of the limit sets of belief-free equilibrium payoffs of the corresponding known-state games; that is, there are BFXE in which players eventually obtain payoffs almost as if they commonly knew the state and played a belief-free equilibrium for that state. This result is not an immediate consequence of individual learning, because even if players have learned the true state from past private signals, they do not know what the opponent has learned in the past play and hence it is not obvious whether they are willing to play an equilibrium of the known-state game.

We begin with introducing notation and informational conditions imposed in this section. *Player i's action plan* is $\vec{\alpha}_i = (\alpha_i^R)_{R\in\mathscr{R}}$ such that $\alpha_i^R \in \triangle R_i$ for each $R \in \mathscr{R}$. In words, an action plan $\vec{\alpha}_i$ specifies what action to play for each public signal $R \in \mathscr{R}$, in such a way that the specified (possibly mixed) action α_i^R is chosen from the recommended set $\triangle R_i$. Let $\vec{A}_i$ denote the set of all such player i's (possibly mixed) action plans $\vec{\alpha}_i$. That is, $\vec{A}_i = \times_{R\in\mathscr{R}} \triangle R_i$.

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 the action profile (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 $\vec{\alpha}_{-i}$, ω, and R, let $\Pi^{\omega,R}_{-i}(\vec{\alpha}_{-i})$ be a matrix with rows $\hat{\pi}^{\omega}_{-i}(a_i, \alpha^R_{-i})$ 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 3. 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 (2010); here we consider the joint distribution of actions and signals, while they consider the distribution of signals.

Definition 4. 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 $\Pi_{-i}^{(\omega,\tilde{\omega}),R}(\vec{\alpha}_{-i})$ 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. Note that statewise full rank does not pose any restriction on speed of learning; it may be that the signal distributions at state ω are close to those at state $\tilde{\omega}$, giving rise to a slow learning process. But it does not pose any problem on our analysis, as we consider patient players. Again the definition of statewise full rank here is slightly different from that of Fudenberg and Yamamoto (2010), 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 $\omega \neq \tilde{\omega}$, there is $\vec{\alpha}_{-i}$ that has statewise full rank for this pair at some regime $R \in \mathscr{R}$.

This condition (SFR) requires that for each pair $(\omega, \tilde{\omega})$, players can statistically distinguish these two states. Note that (SFR) is sufficient for each player to learn the true state in the long run, even if there are more than two possible states. To identify the true state, we ask a player to collect private signals and to perform a statistical inference to distinguish ω and $\tilde{\omega}$ for each possible pair $(\omega, \tilde{\omega})$ with $\omega \neq \tilde{\omega}$. Under (SFR), the true state will be selected in the all relevant statistical tests; e.g., if there were three possible states and the true state were ω_1, then ω_1 would be selected in the statistical test for (ω_1, ω_2) and in the one for (ω_1, ω_3). Therefore, if there is a state which is selected in all statistical tests, then she can conclude that it is the true state. One remark is that while each player can learn the true state under (SFR), players do not share any common information, and hence it is unclear if players can coordinate their play state by state. (In Appendix B, we provide an example of an equilibrium where each player privately learns the state but common learning fails.)

Given ω, let G^{ω} denote the infinitely repeated game where players know that the true state is ω. Consider belief-free equilibria of EHO in this known-state game G^{ω}, and let $E^{\omega,p}(\delta)$ be the payoff set of belief-free equilibria with respect to public randomization p in the game G^{ω} given δ. Corollary 1 of EHO shows that, as for BFXE, the payoff set of belief-free equilibria has a product structure for each p; that is, $E^{\omega,p}(\delta) = \times_{i\in I}[m_i^{\omega,p}(\delta), M_i^{\omega,p}(\delta)]$. Here $M_i^{\omega,p}(\delta)$ and $m_i^{\omega,p}(\delta)$ are the maximum and minimum of player i's payoffs attained by belief-free equilibria with respect to p. Let $M_i^{\omega,p}$ and $m_i^{\omega,p}$ be the limit of $M_i^{\omega,p}(\delta)$ and $m_i^{\omega,p}(\delta)$ as $\delta \to 1$, i.e., $M_i^{\omega,p}$ and $m_i^{\omega,p}$ denote the maximum and minimum of player i's payoffs of belief-free equilibria with respect to p in the limit as $\delta \to 1$. EHO show that we can compute $M_i^{\omega,p}$ and $m_i^{\omega,p}$ by simple formulas. (For completeness, we give these formulas in Appendix C.)

The main result of the paper is:

Proposition 3. *If (IFR) hold, then* $\lim_{\delta\to 1} E^p(\delta) = \times_{\omega\in\Omega} \times_{i\in I} [m_i^{\omega,p}, M_i^{\omega,p}]$ *for each* $p \in \triangle\mathscr{R}$ *such that (i)* $M_i^{\omega,p} > m_i^{\omega,p}$ *for all* i *and* ω *and (ii) for each* i *and* (ω,ω')*, there is* $\vec{\alpha}_{-i}$ *that has statewise full rank for* (ω,ω') *at some regime* R *with* $p(R) > 0$*. Hence, if (IFR) and (SFR) hold and if there is* $p \in \triangle\mathscr{R}$ *such that* $M_i^{\omega,p} > m_i^{\omega,p}$ *for all* i *and* ω*, 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 given public randomization p, the limit set of BFXE payoffs is isomorphic to the set of maps from states to belief-free equilibrium payoffs. (Recall that the set $\times_{i\in I}[m_i^{\omega,p}, M_i^{\omega,p}]$ denotes the limit set of belief-free equilibrium payoffs given public randomization p.) In other words, there are BFXE where players eventually obtain payoffs almost as if they commonly learned the state and played a belief-free equilibrium for that state. Note that this result reduces to Proposition 5 of EHO if $|\Omega| = 1$.

For example, consider the secret price-cutting game in Section 3.1. Assume that if the firms knew the distribution of an aggregate shock η, there would be a belief-free equilibrium where the firms earn payoffs Pareto-dominating a static Nash equilibrium payoff; i.e., we assume that there is p such that $M_i^{\omega,p} > m_i^{\omega,p}$ and $M_i^{\omega,p} > g_i^{\omega}(\alpha^{NE,\omega})$ for all i and ω, where $\alpha^{NE,\omega}$ is a Nash equilibrium of the stage game when ω is known. Then the above theorem says that even though the firms do not know the distribution of η, they can still maintain the same level of collusion under (SFR).

An important question is when such p exists, and a sufficient (though not necessary) condition is that the uncertainty on the distribution of η is "small." To see this, suppose that the monitoring structure at different states are "close" in that $|\pi^{\omega}(\sigma|a) - \pi^{\tilde{\omega}}(\sigma|a)| < \varepsilon$ for all a, σ, ω, and $\tilde{\omega}$ where $\varepsilon > 0$ is a small number. Assume that for some state $\omega^* \in \Omega$, there is a belief-free equilibrium where the firms earn payoffs Pareto-dominating a static Nash equilibrium payoff, that is, assume that for some ω^* and p^*, $M_i^{\omega^*,p^*} > m_i^{\omega^*,p^*}$ and $M_i^{\omega^*,p^*} > g_i^{\omega}(\alpha^{NE,\omega^*})$ for each i. From EHO, we know that $M_i^{\omega,p}$ and $m_i^{\omega,p}$ are continuous with respect to π^{ω} almost everywhere; thus when ε is sufficiently small, generically we have $M_i^{\omega,p^*} > m_i^{\omega,p^*}$ and $M_i^{\omega,p^*} > g_i^{\omega}(\alpha^{NE,\omega})$ for all i and ω, which shows that p^* satisfies the assumption. This shows that if the uncertainty is small, the firms can earn the same profit as in the case with no uncertainty. Note that this is not a trivial result, because typically equilibrium strategies of EHO depend on the fine details of the signal distribution and a belief-free equilibrium at state ω^* is not an equilibrium at state $\omega \neq \omega^*$ even if the uncertainty is small.

To give the intuition behind Proposition 3, let us focus on BFXE where players are indifferent over all actions in any period and any state.[10] In our equilibria, (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, players' incentive compatibility is satisfied, as property (i) assures that each player's play is optimal after every history. Note that in these equilibrium strategies, player i's individual state learning is irrelevant to her own continuation payoffs, and influences player $-i$'s payoffs only. Indeed, it

[10] 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 $R = A$.

follows from (i) that player i cannot obtain better payoffs by changing her action contingently on what she learned from the past history. Therefore, she is willing to use information about the true state in order to give appropriate payoffs to her opponent with no concern about her own payoffs.

The formal proof of Proposition 3 is provided in Appendix D, and it consists of two steps. In the first step, we consider a general environment (i.e., we do not assume (IFR) or (SFR)) and develop an algorithm to compute the limit set of BFXE payoffs, $\lim_{\delta \to 1} E^p(\delta)$. 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.[11] To take into account the effect of this trade-off, we build on the linear programming (LP) technique of Fudenberg and Yamamoto (2010), who characterize the limit payoffs of ex-post equilibria in repeated games with public and unknown monitoring technology.[12] Specifically, for each player i and for each weighting vector $\lambda_i = (\lambda_i^\omega)_{\omega \in \Omega} \in \boldsymbol{R}^{|\Omega|}$, we consider a static LP problem whose objective function is the weighted sum of player i's payoffs at different states, and we demonstrate that the limit set of BFXE payoffs for player i is characterized by solving these LP problem for all weighting vectors λ_i. Here the trade-offs between equilibrium payoffs for different states are determined by LP problems for "cross-state directions" λ_i that have non-zero components on two or more states; roughly, low scores in these LP problems mean more trade-offs

[11]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.

[12]Fudenberg and Levine (1994) proposes a linear programming characterization of the equilibrium payoff set in repeated games with public monitoring, and Fudenberg and Yamamoto (2010) extend it to the case where the monitoring structure is unknown.

between payoffs for different states. See Appendix D.1.2 for details.

Then in the second step of the proof, we apply the algorithm developed in the first step to games that satisfy (IFR) and (SFR). We show that (i) under (SFR), the LP problems for all cross-state directions give sufficiently high scores and hence there is no trade-off between equilibrium payoffs at different states, and (ii) under (IFR), the LP problems for other directions ("single-state directions") reduce to the ones which compute the bounds $M_i^{\omega,p}$ and $m_i^{\omega,p}$ of belief-free equilibrium payoffs of the known-state games. Combining these two, we can conclude that $\lim_{\delta\to 1} E^p(\delta) = \times_{\omega\in\Omega}\times_{i\in I}[m_i^{\omega,p}, M_i^{\omega,p}]$. The proof of (i) is similar to the the one by Fudenberg and Yamamoto (2010), and its intuition is simple; under (SFR), player $-i$ can learn the state in the long run and can eventually use different actions at different states, which means that there is no trade-off between player i's payoffs across states. The proof of (ii) is slightly different from the one by Fudenberg and Yamamoto (2010). The key in our proof is that we define individual full rank using joint distributions of (a_{-i}, σ_{-i}) so that all mixed actions have individual full rank under (IFR). Then as shown in Lemma 6 in Appendix D.2, the result immediately follows. On the other hand, Fudenberg and Yamamoto (2010) define individual full rank using distributions of signals only, and with this definition, some mixed action profiles may not have individual full rank even if all pure action profiles have individual full rank. As a result, they need a more careful analysis in order to prove the counterpart of (ii).

Remark 4. As a corollary of Proposition 3, 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 Revisiting the Examples in Section 3

As an application of Proposition 3, we revisit the public goods game in Section 3.2. We have already shown that there are efficient equilibria in the public goods game if the likelihood ratio condition (1) is satisfied. Now we apply Proposition 3 to this example to show that the likelihood ratio condition (1) is not necessary for the existence of asymptotically efficient equilibria. Specifically, instead of the likelihood ratio condition (1), we assume that for each i, there is an action a_i such that $(\pi_i^{\omega_1}(z_i|a_i,a_{-i}))_{z_i} \neq (\pi_i^{\omega_2}(z_i|a_i,a_{-i}))_{z_i}$ for each a_{-i}; this assures that player i can learn the true state ω from observed signals regardless of the opponent's play. We show that there are asymptotically efficient equilibria under this weaker assumption.

It is easy to see that (SFR) holds under this assumption. Note also that (IFR) is satisfied in this example, as actions are observable. Hence, Proposition 3 applies and the limit set of BFXE payoffs is equal to the product of the belief-free equilibrium payoff sets of the known-state games; in particular, we have $\times_{\omega\in\Omega} \times_{i\in I} [m_i^{\omega,p}, M_i^{\omega,p}] \subseteq \lim_{\delta\to 1} E(\delta)$ for each p. EHO show that when actions are observable, the bounds $M_i^{\omega,p}$ and $m_i^{\omega,p}$ of belief-free equilibrium payoffs are computed by the following simple formulas:[13]

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

and

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

We use these formulas to compute the BFXE payoffs in this example. Consider $p \in \triangle\mathscr{R}$ such that $p(A) = 1$ and $p(R) = 0$ for other R. From (3) and (4), we have $M_i^{p,\omega_1} = M_i^{p,\omega_2} = 3$, $m_i^{p,\omega_1} = 0$, and $m_i^{p,\omega_2} = 1$ for each i. Hence $\times_{i\in I}([0,3]\times[1,3]) \subseteq \lim_{\delta\to 1} E(\delta)$, which implies that there is a BFXE approximating $((3,3),(3,3))$ for sufficiently large δ. That is, efficiency is achieved for sufficiently high δ even if the likelihood ratio condition (1) is not satisfied. Also, it is easy to see that the result extends to the case where the payoff function g_i^{ω} is perturbed; as long as the payoff matrix is a prisoner's dilemma at ω_1 and is a chicken game at ω_2, the payoff vector $g(C,C) = (g_i^{\omega}(C,C))_{(i,\omega)}$ can be approximated by a BFXE.

Likewise, we can apply Proposition 3 to the example in Section 3.3 to show that there are asymptotically efficient equilibria. Recall that in Section 3.3, we construct a BFXE where players learn the state, but its equilibrium payoff is

[13] In words, $M_i^{\omega,p}$ is equal to 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 equal to 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.

bounded away from the efficient payoff $((1,1),(1,1))$. Now we show that there are BFXE approximating the payoff $((1,1),(1,1))$. To do so, note that both (IFR) and (SFR) are satisfied in this example, so that from Proposition 3, we have $\times_{\omega\in\Omega}\times_{i\in I}[m_i^{\omega,p}, M_i^{\omega,p}] \subseteq \lim_{\delta\to 1} E(\delta)$ for each p. Note also that, since actions are observable, $M_i^{\omega,p}$ and $m_i^{\omega,p}$ are computed by (3) and (4), and we have $M_i^{\omega,p} = 1$ and $m_i^{\omega,p} = 0$ for p such that $p(A) = 1$. Combining these two observations, it follows that there is a BFXE which approximates $((1,1),(1,1))$ when δ is large enough; i.e., the efficient outcome $((1,1),(1,1))$ can be approximated by BFXE. Also, as in the public goods game, the same result holds even if we allow more general signal structures; specifically, the likelihood ratio condition $\pi_i^\omega(z_i^\omega|a) = \frac{2}{3}$ is dispensable, and the signal space Z_i is not necessarily binary. All we need here is (SFR), which is satisfied as long as there is an action a_i such that $(\pi_i^{\omega_1}(z_i|a_i,a_{-i}))_{z_i} \neq (\pi_i^{\omega_2}(z_i|a_i,a_{-i}))_{z_i}$ for each a_{-i}.

5.3 Comparison with Fudenberg and Yamamoto (2010)

This paper investigates the effect of uncertainty about the monitoring structure in repeated games with private monitoring. Fudenberg and Yamamoto (2010) study a similar problem in repeated games with public monitoring, that is, they consider the case where players observe public signals in every period. They find a sufficient condition for the folk theorem; i.e., they show that under some informational condition, any feasible and individually rational payoff can be achievable when players are patient. This means that there are equilibria in which players eventually obtain payoffs almost as if they commonly knew the state and played an equilibrium for that state. Their approach and ours are similar in the sense that both look at ex-post equilibria and characterize the limit equilibrium payoffs using linear programming problems. However, our state-learning theorem is not a corollary of Fudenberg and Yamamoto (2010). Indeed, how players learn the state and use that information in BFXE is different from the one in Fudenberg and Yamamoto (2010) in the following sense.

The key in Fudenberg and Yamamoto (2010) is to look at public strategies where players' play depends only on past public signals. This means that players ignore all private information such as the first- or higher-order beliefs about ω; instead, they perform a statistical test about the true state ω using public signals and determine their continuation play. In other words, players form a "publicly observable dummy belief" about the true state ω which depends only on public information, and effectively adjust their play. This allows players to coordinate their play perfectly, and since the publicly observable belief converges to the true state ω, the target payoff can be achieved state by state. Also, players have no incentive to deviate because any unilateral deviation will be statistically detected and hence will be punished in future. Note that the same idea is used in Wiseman (2012), who studies the case where actions are observable and players receive both public and private signals about the true state. He proves the folk theorem by constructing equilibria where players compute a public dummy belief and adjust their continuation play while all private signals are ignored; his constructive proof illustrates the usefulness of a public dummy belief more explicitly than the non-constructive proof of Fudenberg and Yamamoto (2010).

When we consider private monitoring, the above idea does not work because there is no public information; players cannot form a public dummy belief and they need to use private signals to learn the true state. Thus in general, players' higher-order beliefs are relevant to their incentives, which makes the analysis intractable. To avoid such a complication, we consider equilibria where each player makes her opponent indifferent over the relevant actions given any history. This belief-free property assures that players' higher-order beliefs are irrelevant to best replies and incentive compatibility is automatically satisfied.[14] Of course, requiring players to be indifferent comes at a cost in the sense that it is much stronger than sequential rationality; specifically, we need to find a strategy profile which satisfies all the indifference conditions independently of the true state ω. Nonetheless, we find that this requirement still leaves enough strategies so that BFXE can support many non-trivial payoffs (including Pareto-efficient outcomes) if (SFR) is satisfied so that players can individually learn the true state. In other words, our result shows that ex-post equilibria work nicely even if we look at the case where

[14]Indeed, we can formally show that players' higher-order beliefs are irrelevant to the set of BFXE payoffs in the following sense. As shown in Appendix D.1.1, player i's equilibrium payoff set given δ is the largest fixed point of the operator B_i^p, and this operator depends on the signal distribution π only through the marginal distribution π_{-i}. This means that the equilibrium payoff set is the same even if the correlation between private signals changes and players' higher-order beliefs are perturbed. (Note that a change in the correlation influences players' first-order beliefs as well, but the above result shows that such perturbations of first-order beliefs are irrelevant to the equilibrium payoff set.)

players learn the state from private signals (so that they cannot coordinate their play) and even if we impose many indifference conditions.

In addition to the issue discussed above, note that we explicitly construct BFXE in some examples. With these constructions, it is easy to see how players learn the true state from noisy signals in ex-post equilibria. Fudenberg and Yamamoto (2010) do not provide such a result.

6 Conditionally Independent Signals

6.1 BFXE and Review Strategies

In repeated games with private monitoring and with a known state, the set of belief-free equilibrium payoffs is typically a strict subset of feasible and individually rational payoff set. To attain a larger payoff set, several papers combine the idea of review strategies and belief-free equilibria (*belief-free review-strategy equilibria* of Matsushima (2004), EHO, Yamamoto (2007), and Yamamoto (2012)); this approach 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 private signals. After a T-period play, the pooled private signals are used to test whether the opponent deviated or not; then the law of large numbers assures that a player can obtain precise information about the opponent's action from this statistical test. 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 subsection, we show that this approach can be extended to the case where players do not know the true state, although the constructive proof of the existing 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 we compute its equilibrium payoff set. We find that if the signal distribution satisfies some informational conditions, there are sequential equilibria where players eventually obtain payoffs almost as if they commonly knew the state and played a belief-free review-strategy equilibrium for that state. Then in the next subsection, we apply this result to a secret price-cutting game, and show that cartel is self-enforcing even if firms do not have precise information about the market demand. Also we give a simple equilibrium construction.

As mentioned, the past work has shown that review strategies work well for games with independent monitoring.[15] Here 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 $\sigma_0 \in \Sigma_0$.

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).

In addition to (Weak-CI), we assume that the signals distribution has full support.

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

Definition 5. 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.

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 (2012) show, if the signal distribution is weakly conditionally independent, then these values are calculated by the following formulas:

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

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

Note that these formulas are similar to (3) and (4) 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 establishes that if the signal distribution is weakly conditionally independent and if each player can privately learn the true state from observed signal distributions, then there are sequential equilibria where players eventually obtain payoffs almost as if they commonly knew the state and played a belief-free review-strategy equilibrium for that state. Note that this result reduces to Proposition 10 of EHO if $|\Omega| = 1$.

Proposition 4. *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 $\delta \to 1$.*

The proof of this proposition is parallel to that of Proposition 3. Recall that the proof of Proposition 3 consists of two steps; we first develop the linear programming technique to compute the limit set of BFXE payoffs for general environments, and then apply it to games that satisfy the identifiability conditions. Here we follow a similar two-step procedure to prove Proposition 4: We first characterize the limit set of review-strategy equilibrium payoffs for general environments by extending the linear programming technique in Appendix D, and then apply it to games that satisfy the informational conditions. See Appendix E for details.

Remark 5. In Proposition 4, we assume the signal distribution to be weakly conditionally independent. The result here is robust to a perturbation of the signal distribution; that is, any interior point of $\bigcup_{p\in\triangle\mathscr{R}} \times_{i\in I} \times_{\omega\in\Omega} [n_i^{\omega,p}, N_i^{\omega,p}]$ is achieved by a sequential equilibrium if the discount factor is sufficiently close to one and if the signal distribution is sufficiently close to a weakly-conditionally-independent distribution. See Yamamoto (2012) for more details.

6.2 Secret Price-Cutting

Now we apply Proposition 4 to the secret price-cutting game in Section 3.1, and show that firms can maintain a self-enforcing cartel agreement even if they do now know how profitable the market is. To make our analysis simple, suppose that there are only two possible states and $A_i = \{C, D\}$; i.e., in every period, firm i chooses either the high price C or the low price D.

We assume that u_i and π are such that (SFR) and (Weak-CI) hold,[16] and such that the stage game is the prisoner's dilemma for both states; i.e., (C,C) is efficient but D dominates C at each state. Then Proposition 4 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 $N_i^{\omega,p} = g_i^{\omega}(C,C)$ and $n_i^{\omega,p} = g_i^{\omega}(D,D)$ for each i and ω. Therefore the efficient payoff $g(C,C)$ can be approximated by a sequential equilibrium.

Also, in this example, we can explicitly construct asymptotically efficient equilibria. The equilibrium construction here is an extension of the BFXE in Section 3.2. Specifically, the infinite horizon is regarded as a sequence of review phases with T periods, and in each review phase, player i is either in "reward state" $x(1)$ or "punishment state" $x(2)$. When player i is in the reward state $x(1)$, she chooses the high price C for T periods to reward the opponent. On the other hand, when she is in the punishment state $x(2)$, she chooses the low price D for T periods to punish the opponent. At the end of each review phase, player i transits over $x(1)$ and $x(2)$, where the transition probability depends on the recent

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

T-period history.

As in Section 3.2, let $v_2(k) = (v_2^{\omega_1}(k), v_2^{\omega_2}(k))$ denote the target payoff of player 2 when player 1's current state is $x(k)$. Set $v_2^{\omega}(1) = g_2^{\omega}(C,C) - \varepsilon$ and $v_2^{\omega}(2) = g_2^{\omega}(D,D) + \varepsilon$ for each ω where ε is a small positive number; that is, we let the target payoff at the reward state be close to the payoff by (C,C), and the target payoff at the punishment state be close to the payoff by (D,D).

The key of our equilibrium construction is to choose player i's transition rule carefully so that player $-i$ is indifferent between being in $x(1)$ and in $x(2)$ in the initial period of a review phase, regardless of the state of the world ω. For example, suppose that player 1 is in the reward state $x(1)$ and will choose C for the next T periods. Since $g_2^{\omega}(C,D) > g_2^{\omega}(C,C) > v_2^{\omega}(1)$ for each ω, player 2's average payoff for the next T periods will be greater than the target payoff $v_2^{\omega}(1)$ regardless of the true state ω and of what player 2 will do. To offset this extra profit, player 1 will switch to the punishment state $x(2)$ after the T-period play with positive probability. Specifically, at the end of the review phase with length T, player 1 performs statistical tests about the true state ω and about player 2's play using the information pooled within the T periods, and then determines the transition probability. This transition rule is an extension of that in Section 3.2; recall that in the automaton constructed in Section 3.2, the transition probability β depends both on an observed action a_2 and on a private signal z_1 which is sufficiently informative about ω in the sense that the likelihood ratio condition (1) is satisfied. Here in the secret price-cutting model, actions are not directly observable and the likelihood ratio condition may not be satisfied; instead, player 1 aggregates information during T periods to perform statistical tests about a_2 and ω. This allows player 1 to obtain (almost) precise information about a_2 and ω, so that as in Section 3.2, we can find transition probabilities which make player 2 indifferent between being at $x(1)$ and $x(2)$. Also, we can show that when player 1 uses some sophisticated statistical tests, it is suboptimal for player 2 to mix C and D in a T-period play, which means that player 2 is willing to follow the prescribed strategy. The construction of the statistical tests is similar to that in Section 3.2.3 of Yamamoto (2012), and hence omitted.[17]

[17] More specifically, the construction of the statistical test here is very similar to that for the case where the state ω is known and the opponent has four possible actions, because in this example, player i needs to identify a pair (ω, a_{-i}) of the state of the world and the opponent's action and there are four possible pairs (ω, a_{-i}).

The same argument applies to the case where player 1's current state is $x(2)$; we can show that there is a transition rule after $x(2)$ such that player 2 is indifferent being at $x(1)$ and $x(2)$ and is not willing to mix C and D in a T-period play for each state ω, and such that the target payoff $v_2(2)$ is exactly achieved.

We can define player 2's strategy in the same way, and it is easy to see that the pair of these strategies constitute an equilibrium. In particular, when the initial state is $x(1)$ for both players, the equilibrium payoff is $v_i(1)$ for each player i. Since ε can be arbitrarily small, the equilibrium payoff is almost efficient.

7 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 show that 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 commonly learned the true state and played a belief-free equilibrium for that state. Also, we describe equilibrium strategies in some examples, which illustrates how players learn the state and use that information in ex-post equilibria.

As mentioned, 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. Whether the folk theorem obtains by considering belief-based equilibria is an interesting future research.[18]

[18] Throughout this paper, we have assumed that players cannot communicate with each other; so an interesting question is whether a larger payoff set can be attained when we allow players to communicate. For known-state games, Kandori and Matsushima (1998) (as well as Compte (1998)) show that the folk theorem obtains under private monitoring if players can communicate. By combining their proof techniques with the ex-post equilibrium approach of Fudenberg and Yamamoto (2010), we can show that their result extend to the case of unknown monitoring structure; i.e., the folk theorem obtains under mild informational conditions even if the state of the world is unknown. Detailed manuscripts are available upon request.

Appendix A: Equilibrium Strategies in Public Goods Provision

In this appendix, we complete the equilibrium construction in the public goods game in Section 3.2. Our goal is to choose β and γ such that player 2's target payoff $v_2(k)$ is exactly achieved at both $x(1)$ and $x(2)$.

With an abuse of notation, we write $\pi_1^{\omega}(z_1^{\tilde{\omega}}|a)$ for $\pi_1^{\omega}(z_1^{\tilde{\omega}}(a)|a)$; that is, $\pi_1^{\omega}(z_1^{\tilde{\omega}}|a)$ means the probability that player 1 observes $z_1^{\tilde{\omega}}(a)$ given (ω, a). Recall that $\beta(C, z_1) = 0$ for all z_1. We set

$$\beta(D,z_1) = \begin{cases} \dfrac{1-\delta}{\delta} \cdot \dfrac{1 + 2\pi_1^{\omega_2}(z^{\omega_2}|C,D) - 3\pi_1^{\omega_1}(z^{\omega_2}|C,D)}{2\pi_1^{\omega_2}(z^{\omega_2}|C,D) - 2\pi_1^{\omega_1}(z^{\omega_2}|C,D)} & \text{if } z_1 = z_1^{\omega_2}(C,D) \\ \\ \dfrac{1-\delta}{\delta} \cdot \dfrac{2\pi_1^{\omega_2}(z^{\omega_2}|C,D) - 3\pi_1^{\omega_1}(z^{\omega_2}|C,D)}{2\pi_1^{\omega_2}(z^{\omega_2}|C,D) - 2\pi_1^{\omega_1}(z^{\omega_2}|C,D)} & \text{otherwise} \end{cases},$$

$$\gamma(C,z_1) = \begin{cases} \dfrac{1-\delta}{\delta} \cdot \dfrac{1 + 2\pi_1^{\omega_1}(z^{\omega_1}|D,C) - 3\pi_1^{\omega_2}(z^{\omega_1}|D,C)}{\pi_1^{\omega_1}(z^{\omega_1}|D,C) - \pi_1^{\omega_2}(z^{\omega_1}|D,C)} & \text{if } z_1 = z_1^{\omega_1}(D,C) \\ \\ \dfrac{1-\delta}{\delta} \cdot \dfrac{2\pi_1^{\omega_1}(z^{\omega_1}|D,C) - 3\pi_1^{\omega_2}(z^{\omega_1}|D,C)}{\pi_1^{\omega_1}(z^{\omega_1}|D,C) - \pi_1^{\omega_2}(z^{\omega_1}|D,C)} & \text{otherwise} \end{cases},$$

$$\gamma(D,z_1) = \begin{cases} \dfrac{1-\delta}{\delta} \cdot \dfrac{3 + 4\pi_1^{\omega_2}(z^{\omega_2}|D,D) - 7\pi_1^{\omega_1}(z^{\omega_2}|D,D)}{2\pi_1^{\omega_2}(z^{\omega_2})|D,D) - 2\pi_1^{\omega_1}(z^{\omega_2}|D,D)} & \text{if } z_1 = z_1^{\omega_2}(D,D) \\ \\ \dfrac{1-\delta}{\delta} \cdot \dfrac{4\pi_1^{\omega_2}(z^{\omega_2}|D,D) - 7\pi_1^{\omega_1}(z^{\omega_2}|D,D)}{2\pi_1^{\omega_2}(z^{\omega_2}|D,D) - 2\pi_1^{\omega_1}(z^{\omega_2}|D,D)} & \text{otherwise} \end{cases}.$$

Note that β and γ are in the interval $(0,1)$ when δ is large enough. Also we can check that for each ω the following equalities are satisfied:

$$v_2^{\omega}(1) = (1-\delta)g_2^{\omega}(C,C) + \delta \sum_{z_1} \pi_2^{\omega}(z_1|C,C)[\beta(C,z_1)v_2^{\omega}(2) + (1-\beta(C,z_1))v_2^{\omega}(1)],$$

$$v_2^{\omega}(1) = (1-\delta)g_2^{\omega}(C,D) + \delta \sum_{z_1} \pi_2^{\omega}(z_1|C,D)[\beta(D,z_1)v_2^{\omega}(2) + (1-\beta(D,z_1))v_2^{\omega}(1)],$$

$$v_2^{\omega}(2) = (1-\delta)g_2^{\omega}(D,C) + \delta \sum_{z_1} \pi_2^{\omega}(z_1|D,C)[\gamma(C,z_1)v_2^{\omega}(1) + (1-\gamma(C,z_1))v_2^{\omega}(2)],$$

$$v_2^{\omega}(2) = (1-\delta)g_2^{\omega}(D,D) + \delta \sum_{z_1} \pi_2^{\omega}(z_1|D,D)[\gamma(D,z_1)v_2^{\omega}(1) + (1-\gamma(D,z_1))v_2^{\omega}(2)].$$

The first equality shows that when player 1 begins her play with state $x(1)$ and when player 2 chooses C today, then the target payoff $v_2^{\omega}(1)$ is achieved at both ω.

The second equality shows that the same target payoff $v_2^{\omega}(1)$ is still achieved even when player 2 chooses D rather than C. Combining these two, we can conclude that player 2 is indifferent between C and D if player 1 is in state $x(1)$. The next two equalities show that the same is true when player 1 is in state $x(2)$, that is, if player 1's current state is $x(2)$, player 2 is indifferent between C and D and the target payoff $v_2^{\omega}(2)$ is exactly achieved at both ω. So β and γ specified above satisfy all the desired conditions.

Appendix B: Failure of Common Learning

In this appendix, we present an example where players do not achieve approximate common knowledge but players adjust their actions according to their own individual learning and obtain high payoffs state by state. The example here is a simple extension of that in Section 4 of Cripps, Ely, Mailath, and Samuelson (2008).

The following notation is useful. Let Z_i be the set of all non-negative integers, i.e., $Z_i = \{0,1,2,\cdots\}$. Let $Z = \times_{i} \in IZ_i$. In the example of Cripps, Ely, Mailath, and Samuelson (2008), each player i observes a noisy signal $z_i \in Z_i$ about the true state $\theta \in \{\theta', \theta''\}$ in every period. Let $\hat{\pi}^1 \in \triangle Z$ denote the joint distribution of $z = (z_1, z_2)$ at state θ', and let $\hat{\pi}^2 \in \triangle Z$ denote the distribution at state θ'', (For example, the probability of the signal profile $z = (0,0)$ is θ' given $\hat{\pi}^1$, and θ'' given $\hat{\pi}^2$.)

In this appendix, we consider the following example. There are two players and two possible states, so that $\Omega = \{\omega_1, \omega_2\}$. Players have a common initial prior over states, $\frac{1}{2}$-$\frac{1}{2}$. Each player has two possible actions; $A_1 = \{U, D\}$ and $A_2 = \{L, R\}$. Actions are observable, and in addition each player i observes a noisy signal $z_i \in Z_i$ about the true state in every period. The joint distribution of $z = (z_1, z_2)$ is dependent only on the true state (i.e., it does not depend on actions played), and assume that the joint distribution of z is exactly the same as the example of Cripps, Ely, Mailath, and Samuelson (2008); i.e., the joint distribution is equal to $\hat{\pi}^1$ at state ω_1 and to $\hat{\pi}^2$ at state ω_2.[19] The expected payoffs for state

[19] In this setup, player i's signal space is $\Sigma = A \times Z_i$, which is not a finite set. But it is straightforward to see that the results in Section 5 extend to the case of infinitely many signals, by considering a finite partition of Z_i. See Ishii (2009). Also, a version of (SFR) is satisfied in this example.

ω_1 is shown in the left panel, and those for state ω_2 is in the right.

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

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

In this stage game, player 1's action influences player 2's payoff only. Specifically, the action U is efficient (i.e., gives high payoffs to player 2) at state ω_1, while the action D is efficient at state ω_2. Likewise, player 2's action influences player 1's payoff only; the efficient action is L at state ω_1 and is R at state ω_2. Note that players are indifferent between two actions given any state, thus all action profiles are ex-post equilibria of the one-shot game.

Given a natural number T, let $s(T)$ be the following strategy profile of the infinitely repeated game:

- Players mix two actions with $\frac{1}{2}$-$\frac{1}{2}$ in period t for each $t = 1, \cdots, T$.
- Let $q_i(h_i^T|s(T)) \in \triangle\Omega$ be player i's belief about the state at the end of period T. From period $T+1$ on, player 1 chooses U forever if $q_1(h_1^T|s(T))[\omega_1] \geq \frac{1}{2}$, and chooses D forever otherwise. Likewise, player 2 chooses L forever if $q_2(h_2^T|s(T))[\omega_1] \geq \frac{1}{2}$, and chooses R forever otherwise.

In words, players try to learn the true state in the first T periods ("the learning phase"), and then adjust their continuation play to achieve high payoffs state by state. This strategy profile $s(T)$ is a stationary BFXE given any T, since actions do not influence the distribution of z and all action profiles are ex-post equilibria of the one-shot game.

In this example, the limit equilibrium payoff (as $\delta \to 1$) approximates the efficient payoff vector $((1,1),(1,1))$ for T sufficiently large, since each player can obtain arbitrarily precise information about the state during the learning phase. On the other hand, the state ω_2 cannot be (approximate) common knowledge during the learning phase, even if we take T sufficiently large. Indeed, as Section 4 of Cripps, Ely, Mailath, and Samuelson (2008) shows, there is $p > 0$ such that given any T sufficiently large, the state ω_2 can never be common p-belief at date T conditional on the strategy profile $s(T)$.

Appendix C: Computing $M_i^{\omega,p}$ and $m_i^{\omega,p}$

In this appendix, we provide a formula to compute $M_i^{\omega,p}$ and $m_i^{\omega,p}$, the maximum and minimum of belief-free equilibrium payoffs in the limit as $\delta \to 1$. (4) and (5) of EHO show how to compute the maximum and minimum of belief-free equilibrium payoffs. In our notation,

$$M_i^{\omega,p} = \sup_{\vec{\alpha}_{-i}} M_i^{\omega,p}(\vec{\alpha}_{-i}),$$
$$m_i^{\omega,p} = \inf_{\vec{\alpha}_{-i}} m_i^{\omega,p}(\vec{\alpha}_{-i})$$

where

$$M_i^{\omega,p}(\vec{\alpha}_{-i}) = \max_{\substack{v_i^{\omega}\in\mathbb{R} \\ x_i^{\omega}:\mathscr{R}\times A_{-i}\times\Sigma_{-i}\to\mathbb{R}}} v_i^{\omega} \quad \text{subject to}$$

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

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

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

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

$$m_i^{\omega,p}(\vec{\alpha}_{-i}) = \min_{\substack{v_i^{\omega}\in\mathbb{R} \\ x_i^{\omega}:\mathscr{R}\times A_{-i}\times\Sigma_{-i}\to\mathbb{R}}} v_i^{\omega} \quad \text{subject to}$$

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

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

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

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

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

Appendix D: Characterizing the Set of BFXE Payoffs

In this appendix, we prove Proposition 3. Appendix D.1 provides a preliminary result, that is, we consider general environments (i.e., we do not assume (IFR) or (SFR)) and develop an algorithm to compute the set of BFXE payoffs in the limit as $\delta \to 1$. This is an extension of the linear programming techniques of Fudenberg and Levine (1994), EHO, and Fudenberg and Yamamoto (2010). Then in Appendix D.2, we apply the algorithm to games that satisfy (IFR) and (SFR) to prove Proposition 3.

D.1 Linear Programming Problems and BFXE

D.1.1 Individual Ex-Post Generation

To begin, we give a recursive characterization of the set of stationary BFXE payoffs for general discount factor δ. This is a generalization of the self-generation theorems of Abreu, Pearce, and Stacchetti (1990) and EHO.

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 (R,a,σ). Note that player i's continuation payoff $w_i^\omega(y,a,\sigma)$ at state ω does not depend on (a_i,σ_i), since 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 $w_i(y,a_{-i},\sigma_{-i}) = (w_i^\omega(y,a_{-i},\sigma_{-i}))_{\omega\in\Omega}$.

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 6. 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 $\vec{\alpha}_{-i} \in \vec{A}_{-i}$ and a function $w_i : \mathscr{R} \times A_{-i} \times \Sigma_{-i} \to W_i$ such that

$$v_i^\omega = \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^R(a_{-i}) \left[\begin{array}{l} (1-\delta)g_i^\omega(a_i^R, a_{-i}) \\ +\delta\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot w_i^\omega(R, a_{-i}) \end{array} \right] \tag{5}$$

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

$$v_i^\omega \geq \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^R(a_{-i}) \left[\begin{array}{l} (1-\delta)g_i^\omega(a_i^R, a_{-i}) \\ +\delta\pi_{-i}^\omega(a_i^R, a_{-i}) \cdot w_i^\omega(R, a_{-i}) \end{array} \right] \tag{6}$$

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

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 $R_i \subseteq A_i$ 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 7. A subset W_i of $\mathbb{R}^{|\Omega|}$ is *individually ex-post self-generating with respect to* (δ, p) if $W_i \subseteq B_i^p(\delta, W_i)$.

The following two propositions provide a recursive characterization of the set of stationary BFXE payoffs for any discount factor $\delta \in (0,1)$. Proposition 5, 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 6 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. The proofs of the propositions are similar to Abreu, Pearce, and Stacchetti (1990) and EHO, and hence omitted.

Proposition 5. *For every* $\delta \in (0,1)$ *and* $p \in \triangle\mathscr{R}$, $E^p(\delta) = \times_{i\in I} B_i^p(\delta, E_i^p(\delta))$.

Proposition 6. *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* $\times_{i\in I} W_i \subseteq E^p(\delta)$.

D.1.2 Linear Programming Problem and Bound of $E^p(\delta)$

Here 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 at 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 $R \in \mathscr{R}$, $i \in I$, $\delta \in (0,1)$, $\vec{\alpha}_{-i} \in \vec{A}_{-i}$, 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) (5) holds for all $\omega \in \Omega$ and $(a_i^R)_{R\in\mathscr{R}}$ s.t. $a_i^R \in R_i$ for each $R \in \mathscr{R}$,

(ii) (6) 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 $\bar{k} > 0$ there is (v_i, w_i) satisfying all the constraints and $\lambda_i \cdot v_i > \bar{k}$, 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.

Note that constraint (iii) allows "utility transfer across states." To see how this constraint works, recall that player $-i$ obtains (possibly noisy) information about the true state from her private signal σ_{-i}. Let λ_i be such that $\lambda_i^{\omega} > 0$ for all ω to make our exposition as simple as possible. Constraint (iii) makes the following scheme feasible:

- If player $-i$ observes a signal σ_{-i} which indicates that the true state is likely to be ω, then she chooses a continuation strategy (i.e., choose a continuation payoff vector $w_i(R,a_{-i},\sigma_{-i})$) that yields higher payoffs to player i at state ω but lower payoffs at state $\tilde{\omega}$.

- If player $-i$ observes a signal $\tilde{\sigma}_{-i}$ which indicates that the true state is likely to be $\tilde{\omega}$, then she chooses a continuation strategy that yields higher payoffs to player i at state $\tilde{\omega}$ but lower payoffs at state ω.

In this scheme, player $-i$ adjusts her continuation play contingently on her state learning, so that high expected continuation payoffs are obtained at both states. This shows that under constraint (iii), state learning can help improving players' overall payoffs. Note that this issue does 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) $$\max_{\substack{v_i \in \mathbb{R}^{|\Omega|} \\ x_i:\mathscr{R}\times A_{-i}\times\Sigma_{-i}\to\mathbb{R}^{|\Omega|}}} \lambda_i \cdot v_i \qquad \text{subject to}$$

(i) $$v_i^{\omega} = \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R}(a_{-i}) \begin{bmatrix} g_i^{\omega}(a_i^R,a_{-i}) \\ +\pi_{-i}^{\omega}(a_i^R,a_{-i}) \cdot x_i^{\omega}(R,a_{-i}) \end{bmatrix}$$

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

(ii) $$v_i^{\omega} \geq \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^{R}(a_{-i}) \begin{bmatrix} g_i^{\omega}(a_i^R,a_{-i}) \\ +\pi_{-i}^{\omega}(a_i^R,a_{-i}) \cdot x_i^{\omega}(R,a_{-i}) \end{bmatrix}$$

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

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

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 finding 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 α_{-i}^R, and the remaining components $\alpha_{-i}^{\tilde{R}}$ for $\tilde{R} \neq R$ are irrelevant. Therefore, we have

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

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

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

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),

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

for all $(a_i^R)_{R\in\mathscr{R}}$ such that $a_i^R \in R_i$ 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 D.1.4 for completeness.

D.1.3 Computing $E(\delta)$ with Patient Players

In Appendix D.1.2, 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 8. 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 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 9. 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.[20]

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 D.1.4 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.

[20] 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 $\mathbb{R}^{|\Omega|}$.

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 7. *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 limit set of all stationary BFXE payoffs, $E(\delta) = \bigcup_{p\in\triangle\mathscr{R}} E^p(\delta)$. This is a counterpart of Proposition 4 of EHO.

Proposition 8. *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 7, 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 $i \in I$.

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.*

D.1.4 Proofs of Lemmas 2 and 4

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, $E_i^p(\delta) \subseteq coE_i^p(\delta) \subseteq Q_i^p$.*

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

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 $\delta \in (\overline{\delta}, 1)$, $\times_{i\in I} W_i \subseteq E^p(\delta)$.*

Proof. From lemma 1(c), Q_i^p 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 $\text{int}W_i$ 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 $\text{int}W_i$ 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.*

D.2 Proof of Proposition 3

Proposition 3. *If (IFR) hold, then* $\lim_{\delta\to 1} E^p(\delta) = \times_{\omega\in\Omega} \times_{i\in I} [m_i^{\omega,p}, M_i^{\omega,p}]$ *for each* $p \in \triangle\mathscr{R}$ *such that (i)* $M_i^{\omega,p} > m_i^{\omega,p}$ *for all* i *and* ω *and (ii) for each* i *and* (ω,ω')*, there is* $\vec{\alpha}_{-i}$ *that has statewise full rank for* (ω,ω') *at some regime* R *with* $p(R) > 0$*. Hence, if (IFR) and (SFR) hold and if there is* $p \in \triangle\mathscr{R}$ *such that* $M_i^{\omega,p} > m_i^{\omega,p}$ *for all* i *and* ω*, then* $\lim_{\delta\to 1} E(\delta) = \bigcup_{p\in\triangle\mathscr{R}} \times_{\omega\in\Omega} \times_{i\in I} [m_i^{\omega,p}, M_i^{\omega,p}]$.

Proposition 7 in Appendix D.1.3 shows that the limit equilibrium payoff set is characterized by a series of linear programming problems (LP-Individual). To prove Proposition 3, we compute the maximal score of (LP-Individual) for each direction λ_i for games that satisfy (IFR) and (SFR).

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 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* $\lambda_i^{\tilde{\omega}} \neq 0$*, we have* $k_i^p(\vec{\alpha}_{-i}, \lambda_i) = \infty$.

This lemma is analogous to Lemma 6 of Fudenberg and Yamamoto (2010), and we give the formal proof in Appendix D.2.1 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.

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 $k_i^p(\lambda_i) = M_i^{\omega,p}$. (See Appendix E.) 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 6. *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 3; we use Lemmas 5 through 6 to compute the scores of (LP-Individual) for various directions.

Proof of Proposition 3. From Proposition 8 of Appendix D.1.3, 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 Lemma 5 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 $\lambda_i \notin \Lambda_i^*$, Therefore, $Q_i^p = \bigcap_{\lambda_i\in\Lambda} H_i^p(\lambda_i) = \bigcap_{\lambda_i\in\Lambda^*} H_i^p(\lambda_i)$. Note that, from Lemma 6, 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, $Q_i^p = \bigcap_{\lambda_i\in\Lambda^*} H_i^p(\lambda_i) = \times_{\omega\in\Omega}[m_i^{\omega,p}, M_i^{\omega,p}]$, and Propositions 7 and 8 apply. *Q.E.D.*

D.2.1 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 $\lambda_i^{\tilde{\omega}} \neq 0$, we have $k_i^p(\vec{\alpha}_{-i}, \lambda_i) = \infty$.*

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

$$\sum_{a_{-i}\in A_{-i}} \alpha_{-i}^R(a_{-i})\pi_{-i}^{\omega}(a) \cdot z_i^{\omega}(R, a_{-i}) = \frac{\bar{k}}{\delta p(R)\lambda_i^{\omega}} \tag{7}$$

for all $a_i \in A_i$,

$$\sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{R}(a_{-i}) \pi_{-i}^{\tilde{\omega}}(a) \cdot z_i^{\tilde{\omega}}(R, a_{-i}) = 0 \tag{8}$$

for all $a_i \in A_i$, and

$$\lambda_i^{\omega} z_i^{\omega}(R, a_{-i}, \sigma_{-i}) + \lambda_i^{\tilde{\omega}} z_i^{\tilde{\omega}}(R, a_{-i}, \sigma_{-i}) = 0 \tag{9}$$

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 (9) 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

$$\sum_{a_{-i} \in A_{-i}} \alpha_{-i}^{\hat{R}}(a_{-i}) \left[(1-\delta) g_i^{\hat{\omega}}(a) + \delta \pi_{-i}^{\hat{\omega}}(a) \cdot \tilde{w}_i^{\hat{\omega}}(\hat{R}, a_{-i}) \right] = 0 \tag{10}$$

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 $\bar{k} > \max_{(\hat{R}, a_{-i}, \sigma_{-i})} \lambda_i \cdot \tilde{w}_i(\hat{R}, a_{-i}, \sigma_{-i})$, 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 (7) through (9). 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

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

We claim that this (v_i, w_i) satisfies constraints (i) through (iii) in the LP problem. It follows from (10) that constraints (i) and (ii) are satisfied for all $\hat{\omega} \neq \omega, \tilde{\omega}$.

Also, using (7) and (10), we obtain

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

for all $a_i \in A_i$. This shows that (v_i, w_i) satisfies constraints (i) and (ii) for ω. Likewise, from (8) and (10), (v_i, w_i) satisfies constraints (i) and (ii) for $\tilde{\omega}$. Furthermore, using (9) 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.*

Appendix E: Characterizing the Set of Review-Strategy Payoffs

In this appendix, we prove Proposition 4. Appendix E.1 gives a preliminary result; we consider general environments and develop an algorithm to compute the set of review-strategy equilibrium payoffs. Then in Appendix E.2, we apply the algorithm to games that satisfy (IFR), (SFR), and (Weak-CI) and prove Proposition 4.

E.1 Linear Programming Problems and Review Strategies

Here we consider T-period 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. Specifically, we extend the static LP problem of

Appendix D to T-period LP problems, and establish that the intersection of the corresponding hyperplanes is the limit 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 6 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.

In Appendix D, 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. Here 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

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

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:

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

$$\text{(i)} \quad v_i^\omega = \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^R(a_{-i}) \begin{bmatrix} g_i^\omega(a_i^R,a_{-i}) \\ +\pi_{-i}^{T,\omega}(a_i^R,a_{-i})\cdot x_i^\omega(R,a_{-i}) \end{bmatrix}$$

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

$$\text{(ii)} \quad v_i^\omega \geq \sum_{R\in\mathscr{R}} p(R) \sum_{a_{-i}\in A_{-i}} \alpha_{-i}^R(a_{-i}) \begin{bmatrix} g_i^{T,\omega}(s_i^{T,R},a_{-i},\delta) \\ +\pi_{-i}^{T,\omega}(s_i^{T,R},a_{-i})\cdot x_i^\omega(R,a_{-i}) \end{bmatrix}$$

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

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

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

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

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

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 6 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 nonconstant 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.[21]

Let

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

and

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

Note that $k_i^p(T, \lambda_i, K)$ here 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.

The next proposition is a counterpart to Lemma 4, which 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. The proof of the proposition is given in Appendix E.1.1.

Proposition 4. *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 $\delta \to 1$.*

In the proof of the proposition, 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

[21]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.

optimal). That is, here we consider *periodically belief-free* and *periodically ex-post* equilibria.[22] 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 6. 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 (2011b) 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 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 9 satisfies constraints (i) through (iv) of (T-LP) but does not satisfy the uniform incentive compatibility of Kandori and Matsushima (1998).

Remark 7. 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 4.

[22] 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 (2012); 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}.

E.1.1 Proof of Proposition 4

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 10. 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 $\vec{\alpha}_{-i} \in \vec{A}_{-i}$ and a function $w_i : \mathscr{R} \times A_{-i} \times (\Sigma_{-i})^T \to W_i$ such that

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

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

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

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 $W_i \subseteq B_i^p(T, \delta, W_i, p)$

Lemma 6. *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 6. *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 5, where this set is labeled "G.")

Definition 11. 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 $(R, a_{-i}, \sigma_{-i}^1, \cdots, \sigma_{-i}^T)$.

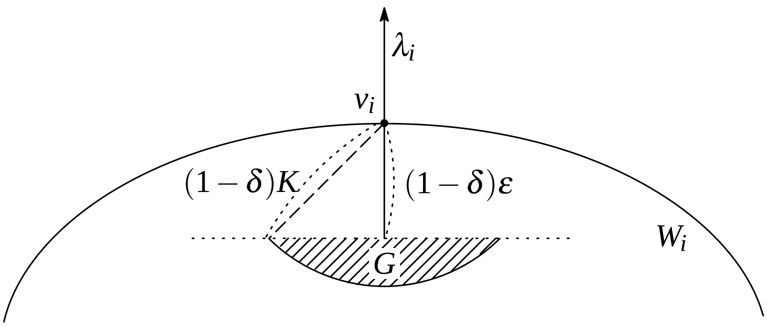

Figure 5: Set G.

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 $\delta \in (\overline{\delta},1)$.*

Proof. Analogous to Fudenberg and Yamamoto (2011b). *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 $\lambda_i \cdot x_{i,\lambda_i,\delta}(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i}) \leq 0$ 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 $\lambda_i \cdot w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i}) \leq \lambda_i \cdot v_i - \frac{1-\delta^T}{\delta^T}\tilde{\varepsilon} < \lambda_i \cdot v_i - (1-\delta^T)\tilde{\varepsilon}$. Also, $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i})$ is within $\frac{1-\delta^T}{\delta^T}\tilde{K}_{\lambda_i}$ of v_i, as $|x_{i,\lambda_i,\delta}(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i})| < K_{\lambda_i}$.)

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 $v_i \in W_i$, $(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i})$, and $\lambda'_i \in \Lambda_i \cap U_{\lambda_i,\delta,v_i}$. (See Figure 6, 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 $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i}) \in G_{v_i,\lambda'_i,\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$ for any $v_i \in W_i$, $(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i})$, and $\lambda'_i \in \Lambda_i \cap U_{\lambda_i,\delta,v_i}$, since $w_{i,\lambda_i,\delta,v_i}(R, a_{-i}, \sigma^1_{-i}, \cdots, \sigma^T_{-i}) \in G_{v_i,\lambda_i,2\varepsilon,\tilde{K}_{\lambda_i},\delta^T}$.

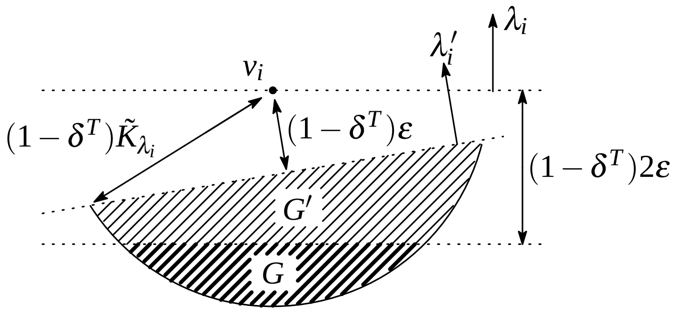

Figure 6: $G \subseteq G'$.

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 4 follows.

E.2 Proof of Proposition 4

Proposition 4. *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 $\delta \to 1$.*

Proposition 4 in Appendix E.1 establishes that the limit set of review-strategy equilibrium payoffs is characterized by a series of linear programming problems (T-LP). To prove Proposition 4, we solve these (T-LP) for various directions for games that satisfy (IFR) and (SFR) and apply Proposition 4. The next lemma is an extension of Lemma 5, which assert that under (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 the lemma does not require the signal distribution to be weakly conditionally independent. The proof of the lemma is found in Appendix E.2.1

Lemma 9. *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 $\lambda_i^{\tilde{\omega}}\neq 0$.*

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. The proof is found in Appendix E.2.1.

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 $\lambda_i^{\omega}=-1$.*

Combining the above three lemmas with Proposition 4, we obtain Proposition 4.

E.2.1 Proofs of Lemmas 9 and 10

Lemma 9. *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 $\lambda_i^{\tilde{\omega}}\neq 0$.*

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,

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

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.*

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

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

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 (2012) 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 (2012) 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.*

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.

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

Cripps, M., J. Ely, G.J. 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.

Dekel, E., D. Fudenberg, and D.K. Levine (2004): "Learning to Play Bayesian Games," *Games and Economic Behavior* 46, 282-303.

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 (2011): "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 (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 (2010): "Repeated Games where the Payoffs and Monitoring Structure are Unknown," *Econometrica* 78, 1673-1710.

Fudenberg, D., and Y. Yamamoto (2011a): "Learning from Private Information in Noisy Repeated Games," *Journal of Economic Theory* 146, 1733-1769.

Fudenberg, D., and Y. Yamamoto (2011b): "The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring," *Journal of Economic Theory* 146, 1664-1683.

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 (2011): "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices," *American Economic Review* 101, 2425-2449.

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 (2011): "Belief-Free Equilibria in Games with Incomplete Information: Characterization and Existence," *Journal of Economic Theory* 146, 1770-1795.

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. (2011): "Weakly Belief-Free Equilibria in Repeated Games with Private Monitoring," *Econometrica* 79, 877-892.

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.

Koren, G. (1992): "Two-Person Repeated Games where Players Know Their Own Payoffs," mimeo.

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 (2011): "Folk Theorems with Bounded Recall and (Almost) Perfect Monitoring," *Games and Economic Behavior* 71, 174-192.

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. (2012): "Robust collusion with private information," *Review of Economic Studies* 79, 778-811.

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

Park, J-H. (2011): "Enforcing International Trade Agreements with Imperfect Private Monitoring," *Review of Economic Studies* 78, 1102-1134.

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.

Rotemberg, J.J., and G. Saloner (1986): "A Supergame-Theoretic Model of Price Wars during Booms," *American Economic Review* 76, 390-407.

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

Shalev, J. (1994): "Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs," *Games and Economic Behavior* 7, 246-259.

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

Sugaya, T. (2010b): "Folk Theorem in a Prisoners' Dilemma without Conditional Independence," 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.

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

Wiseman, T. (2012) “A Partial Folk Theorem for Games with Private Learning,” *Theoretical Economics* 7, 217-239.

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. (2012): “Characterizing Belief-Free Review-Strategy Equilibrium Payoffs under Conditional Independence,” *Journal of Economic Theory* 147, 1998-2027.