

Belief-Based Strategies in the Repeated Prisoners' Dilemma with Asymmetric Private Monitoring*

Bo Chen[†]

Department of Economics
University of Wisconsin - Madison

First Version: July, 27, 2006

This Version: March 29, 2007

Abstract

The "belief-based" approach studies an important class of strategies for repeated games with private monitoring where at each point of the game, each player's optimal continuation strategy is determined by the player's beliefs of the private state of the opponents. This paper extends the "belief-based" approach to the repeated prisoners' dilemma with asymmetric private monitoring technologies. We first find that the previous type of construction in Sekiguchi (1997) and Bhaskar and Obara (2002) may not be sufficient to accommodate all asymmetric private monitoring scenarios, especially when players' private monitoring technologies are sufficiently different. We then modify the previous belief-based strategies by letting the player with smaller observation errors always randomize between "cooperate" and "defect" along the cooperative path of the play. It is shown that full efficiency can be approximated using a modified belief-based strategy profile, provided that observation errors are small and a public randomization device is available. We further construct a complete example to show that the modified "belief-based" strategies can be potentially generalized to other two-player repeated games with almost-perfect private monitoring structures.

1 Introduction

Models of repeated games have been used in many economic applications to show that myopic behavior can be deterred through repetition. However, for players to construct effective deterrents of myopic defections, it is crucial that the players *commonly* observe the history of play. Indeed, many previous works on repeated games have assumed perfect or imperfect public monitoring and have mainly analyzed perfect (public) equilibria of repeated games.¹

*I am grateful to Larry Samuelson for his invaluable advice and guidance in the process of writing this paper. I would also like to thank Jo Hertel, Bill Sandholm, Satoru Takahashi and Yuichi Yamamoto for helpful discussions, and Tanapong Potipiti, Jeremy Sandford for comments. Remaining errors are mine.

[†]1180 Observatory Drive, Madison, Wisconsin 53706 . *Email:* bochen@wisc.edu. *Tel.:* 608-263-3866; *Fax:* 608-262-2033. *Webpage:* <http://www.ssc.wisc.edu/~bchen>.

¹With public monitoring, repeated games admit a recursive structure and equilibrium payoffs can be characterized using dynamic programming techniques (see Fudenberg and Maskin [15] for the perfect monitoring case; Abreu, Pearce and Stachetti [2], and Fudenberg, Levine and Maskin [14] for the imperfect public monitoring case).

Repeated games with imperfect private monitoring where players' actions are private and each player only receives noisy private information on opponents' previous actions in each period are, however, natural settings for many applications as well. A leading example is Stigler's "secret price-cut" model: a repeated price-setting oligopoly model where each firm *only* observes its *own* sales level in each period, which depends on both the prices of the firms and unobservable demand shocks. When a firm's sales level is low, the firm does not know for sure whether it came from other firms' secret price-cutting, or from low market demand².

Since there is no nontrivial public history on which players can coordinate their punishments, repeated games with private monitoring are difficult to analyze³. Nevertheless, interesting results on repeated games with private monitoring have recently been discovered. Naturally, this line of research started from scenarios where the problem caused by private monitoring is minimal. These include, for instance, conditionally independent private monitoring where each player's signals convey no information on others' private signals (Matsushima (2004)); almost-public monitoring where players' private signals are highly correlated for all action profiles played (Mailath and Morris (2002, 2006)); and, more commonly, almost perfect private monitoring where each player's signals are private, but reasonably accurate about opponents' previous actions (Bhaskar and Obara (2002), Ely and Välimäki (2002), Ely, Hörner and Olszewski (2005), Hörner and Olszewski (2006), Piccione (2002), Sekiguchi (1997), Yamamoto (2006)).

In general, the private monitoring literature has put forward two main approaches: the "belief-based" approach and the "belief-free" approach. The belief-free approach was first introduced in the context of a two-player repeated prisoners' dilemma with almost perfect private monitoring by Piccione (2002) and Ely and Välimäki (2002). This approach involves finding a set of continuation strategies with the property that given any feasible belief of the opponent(s), *any* continuation strategy in the set is optimal for a player. Hence, players' beliefs are irrelevant for checking optimality in this approach. The irrelevance of players' beliefs amounts to an enormous simplification of analyzing repeated games with private monitoring and as a result, generalizations of belief-free equilibria to more abstract two-player repeated games have been obtained in the literature.

The belief-based approach was first introduced by Sekiguchi (1997) and further generalized by Bhaskar and Obara (2002)⁴ to repeated prisoners' dilemmas with almost-perfect and *symmetric* private monitoring. The main idea of the belief-based approach is to find a *closed* set of continuation strategies for each player such that each player's optimal continuation strategy is always in the set and a player's optimal continuation strategy *does* depend on the player's beliefs of her opponents' continuation strategies. This paper contributes to the belief-based approach by extending the previous belief-based construction to repeated prisoners' dilemmas with *asymmetric* private monitoring and to other repeated games with private monitoring.

There are several reasons why studying belief-based strategies for asymmetric private monitoring settings is important. First, compared to belief-free strategies⁵, belief-based strategies usually have a clear *coordination interpretation*⁶ and are more robust to perturbations of payoff shocks (Harsanyi (1973)). These properties make belief-based strategies an important class of

²Other prominent examples include subjective evaluations of employees' performance by employers, exchanges of goods with uncertain quality etc.

³Kandori (2002) discusses the main difficulties in the analysis of repeated games with private monitoring.

⁴The equilibria for repeated games with almost-public private monitoring constructed in Mailath and Morris (2002, 2006) are also belief-based.

⁵We compare belief-free strategies and belief-based strategies in more detail in **Section 4**.

⁶That is, a clear link between *current incentives* and *future behavior*.

strategies to study for repeated games with private monitoring.

Secondly, previous belief-based approach has only focused on repeated prisoners' dilemmas with *symmetric* imperfect private monitoring structures. However, in real applications, players/firms often differ in their monitoring technologies (for example, consider Stigler's "secret price-cut" story again: as firms' products are *differentiated*, firms typically face *different* unobservable demand shocks). It is therefore natural to consider the setting of asymmetric private monitoring.

We start our investigation by first constructing belief-based strategies as in Bhaskar and Obara (2002) (we call these *conventional belief-based strategies*) for repeated prisoners' dilemmas with asymmetric private monitoring. We find that conventional belief-based strategies are not sufficient to accommodate private monitoring scenarios, in which players' monitoring technologies are sufficiently asymmetric. The problem is that when players' monitoring technologies are different enough, after a long history of cooperation, the player with a larger observation error does not want to punish a *bad* private signal which indicates the opponent has defected.

Motivated by the above finding, we introduce new belief-based strategies (we call these "*keep them guessing*" strategies) in which the player with smaller observation errors always randomizes between "cooperate" and "defect" along the cooperative path. By introducing small amounts of uncertainty, the player ensures that her opponent never becomes *too sure* about her being in the cooperative state, and so maintains the opponent's incentives to punish bad private signals after a long cooperative private history (of the form "*Cc, ..., Cc*"). We show that the modified belief-based strategy profile constitutes a sequential equilibrium when observation errors are small. In addition, players obtain approximately efficient payoffs in the sequential equilibrium when observation errors become arbitrarily small.

We also show that unlike conventional belief-based strategies, which have (so far) mainly been used to study the repeated prisoners' dilemma with almost-perfect private monitoring, our "*keep them guessing*" strategies can be potentially extended to more general repeated games. We construct non-trivial sequential equilibria for an explicit repeated game (with a 3×3 stage game) to outline the main idea of possible generalizations.

The remainder of the paper is organized as follows. In **Section 2**, we adopt the belief-based approach to analyze a repeated prisoners' dilemma under asymmetric private monitoring. In **Section 3**, we construct an example to show potential extensions of "*keep them guessing*" strategies to general two-player repeated games. Finally, we conclude in **Section 4**. All mathematical details and proofs can be found in Appendix.

2 The Repeated Prisoners' Dilemma with Asymmetric Private Monitoring

In this section, we study belief-based strategies in an infinitely repeated prisoners' dilemma with an asymmetric private monitoring structure.

Consider a two-player (player h and player l) infinitely repeated prisoners' dilemma where at each stage, both players choose actions from $A = \{C, D\}$. The stage game \mathbf{G} with *ex ante* payoffs is described in **Figure 1**. We assume that $G, L > 0$ and $G - L < 1$, so that the stage game is a standard prisoners' dilemma where (C, C) is the (unique) efficient action profile in each stage. Players discount future payoffs at a *common* rate δ and maximize the average discounted sum of payoffs. We also assume that a public randomization device, which is uniformly distributed

over $[0, 1]$, is available for the players.

	C	D
C	1, 1	$-L, 1 + G$
D	$1 + G, -L$	0, 0

FIGURE 1. Stage Game G.

The private monitoring structure is defined as follows: while each player cannot perfectly observe her opponent's actions, she can observe an informative private signal $S_i \in \{c, d\}$ ⁷ ($i \in \{l, h\}$) at the end of each stage, where signal c is more likely when the opponent has played C , and signal d is more likely when the opponent has played D . Different from Sekiguchi (1997) and Bhaskar and Obara (2002), we assume the private monitoring structure to be *asymmetric*. That is, given *any* action profile (a_l, a_h) , the probability that *only* player h (resp. l) receives a wrong signal is p_h (resp. p_l), while the probability of both players seeing wrong signals is q (we assume that $\frac{1}{2} > p_h + q > p_h > p_l > 0$). Note that we have *not* imposed any restriction on the relationship between q and (p_h, p_l) , hence the private monitoring structure accommodates both the case of *conditionally independent* private signals and the case of *correlated* private signals. In addition, the strategic form of the stage game is assumed to be fixed and does not vary for different values of p_h, p_l , and q .⁸

In the remainder of this section, we first follow Bhaskar and Obara (2002) to construct a *conventional* belief-based strategy profile for the asymmetric private monitoring setting. We show that the constructed strategy profile is not a Nash equilibrium when the private monitoring is asymmetric enough. We then modify the conventional belief-based strategies and show that the modified belief-based strategy profile constitutes a sequential equilibrium and the efficient payoff pair $(1, 1)$ can be approximated.

2.1 Conventional Belief-Based Strategies for Repeated Prisoners' Dilemma with Asymmetric Private Monitoring

Before we construct conventional belief-based strategies for the asymmetric private monitoring setting, it is useful to first give a brief description on how full cooperation can be approximated in Bhaskar and Obara (2002). The basic idea of the conventional belief-based approach, which builds on Bhaskar and van Damme (2002), is to use initial randomizations to construct nontrivial sequential equilibria for the repeated prisoners' dilemma with almost-perfect private monitoring.

First, it is well known that with full-support *conditionally independent* private monitoring, a pure grim trigger strategy⁹ profile cannot be an equilibrium in the repeated prisoners' dilemma, no matter how small observation errors are. The reason is, if a player sees a bad signal " d " at the end of *the first period*, the player attributes " d " to noise since she believes that her opponent has cooperated with probability 1. As the observation errors are small and her signal " d " gives her

⁷This (S_i only contains two private signals) is assumed *W.L.O.G.* If a player faces a finite signal space and the private monitoring is almost-perfect, it is possible to rearrange the player's finite private signals into two groups with one group containing *good signals* and the other containing *bad signals*.

⁸This assumption is also maintained in Bhaskar and Obara (2002) where *ex ante* payoffs in each stage are fixed. This is different from that in Sekiguchi (1997) where the *realized* payoffs in each stage are assumed to be fixed. Note that if one *does* follow Sekiguchi (1997) to assume instead that the realized payoffs are fixed, then the payoffs of the stage game will be *asymmetric* as well under asymmetric private signals.

⁹In a pure grim trigger strategy, a player plays " C " in every period on the equilibrium path and a permanent defection is triggered *only* by a single bad private signal " d ".

no information on the opponent's first-period action, the player's belief that her opponent has received "c" and thus will continue to cooperate is almost 1. Hence, it is not optimal for her to respond to the bad signal in the next period. Cooperation collapses because of this unwillingness to punish.

Randomization in the first period ensures that players private signals at the end of the first period convey information on the opponent's previous action and thus can resolve the problem mentioned above (as now, given a player's private history "Cd", it could either be that "d" is an observation error or the opponent has played "D" in the first period and hence "d" is a correct signal). Both Bhaskar and Obara (2002) and Sekiguchi (1997) construct a mixed strategy equilibrium where players randomize between "grim-trigger" strategy and "always-defect" strategy *only* in the first period. There are then mainly two classes of histories to worry about on the equilibrium path: "Cc, Cc, ..., Cc" and "Cc, ..., Cc, Cd". In the first class, players will continue to cooperate as long as observation errors are small enough. In the second, a player with a history "Cc, ..., Cc, Cd" will defect. The reason is that the two *most likely events* (each of these two events involves only *one* observation error) in this case are (1) "d" is the player's own observation error and (2) the opponent has observed "d" in the last period and has switched to the punishment phase. As the two events have approximately equal probabilities, the player is willing to defect as long as she is not very patient.

We now follow Bhaskar and Obara (2002) to construct conventional belief-based strategies for the players under the asymmetric private monitoring setting. Consider the following *partial* continuation strategies¹⁰ for the players ($i \in \{l, h\}$):

$$\begin{aligned} \sigma_i^D &: \begin{cases} \text{play } D, & \text{if } t = 0, \\ \text{play } \sigma_i^D & \text{after "Dc" or "Dd" if } t > 0. \end{cases} \\ \sigma_i^C &: \begin{cases} \text{play } C, & \text{if } t = 0 \\ \text{play } \sigma_i^C & \text{after "Cc", while } \sigma_i^D \text{ after "Cd", if } t > 0. \end{cases} \end{aligned}$$

As in Bhaskar and Obara (2002), we use randomizations in the first period to allow players' behavior to depend on private histories. Specifically, to construct initial randomizations, we present the following "coordination game"¹¹ at $t = 0$ if player i only play " σ_i^C " or " σ_i^D " in the game (note, for example, that V_l^{CC} is player l 's expected continuation value when players play (σ_l^C, σ_h^C)):

	σ_h^C	σ_h^D
σ_l^C	V_l^{CC}, V_h^{CC}	V_l^{CD}, V_h^{DC}
σ_l^D	V_l^{DC}, V_h^{CD}	$0, 0$

FIGURE 2. The "coordination game" at $t = 0$.

Define $\hat{\sigma}$ to be the partial mixed strategy "equilibrium" of the above "coordination game" (players *only* randomize at $t = 0$):

$$\hat{\sigma} = \{\hat{\pi}_l \circ \sigma_l^C + (1 - \hat{\pi}_l) \circ \sigma_l^D, \hat{\pi}_h \circ \sigma_h^C + (1 - \hat{\pi}_h) \circ \sigma_h^D\}, \text{ where}$$

$$\begin{cases} \hat{\pi}_l = \frac{-V_h^{CD}}{V_h^{CC} - V_h^{CD} - V_h^{DC}}; \\ \hat{\pi}_h = \frac{-V_l^{CD}}{V_l^{CC} - V_l^{CD} - V_l^{DC}}. \end{cases} \quad (1)$$

¹⁰A *partial* continuation strategy only specifies what a player should play on the equilibrium path, but *not* off the equilibrium path. Note that σ_i^D and σ_i^C specify the same behavior as "always-defect" and "grim-trigger" specify, respectively, on the equilibrium path.

¹¹Note, however, that (σ_l^C, σ_h^C) is *not* a Nash equilibrium of the asymmetric private monitoring game, for reasons described before.

Now, as in Bhaskar and Obara (2002), the *full* belief-based strategy (where each player's continuation strategy is fully specified at *each* information set) is defined as the following: *player i plays according to the pair* $(\rho_i, \hat{\pi}_j)$, *where* $j \neq i$ and¹²

$$\rho_i : [0, \chi_i^*] \longrightarrow \{C, D, \hat{\pi}_i \circ C + (1 - \hat{\pi}_i) \circ D\}, \text{ s.t. } \rho_i(\mu_i) = \begin{cases} C, & \text{if } \mu_i > \hat{\pi}_j; \\ D, & \text{if } \mu_i < \hat{\pi}_j; \\ \hat{\pi}_i \circ C + (1 - \hat{\pi}_i) \circ D, & \text{if } \mu_i = \hat{\pi}_j. \end{cases} \quad (2)$$

2.2 The Conventional Belief-Based Construction is Not Sufficient to Accommodate All Asymmetric Private Monitoring Settings

We now show that the conventional belief-based construction (in Bhaskar and Obara (2002)) does not automatically carry over to all asymmetric private monitoring scenarios, especially when the private monitoring structure is sufficiently asymmetric. Since this "negative" result¹³ is concerned with a certain class of (private) histories with positive probability *on the equilibrium path*, it is thus relevant to both the construction of belief-based strategies in Bhaskar and Obara (2002) and the path-dominance construction in Sekiguchi (1997).

Specifically, the result of this section is that in a repeated prisoners' dilemma with *sufficiently* asymmetric private monitoring, the type of belief-based strategy profile, $((\rho_l, \hat{\pi}_h), (\rho_h, \hat{\pi}_l))$ defined in (2) above, is *not* a Nash equilibrium: player h may not want to punish player l after a private history of the form " Cc, \dots, Cc, Cd ", which happens with positive probability on the equilibrium path. The intuition is the following: when players adopt the conventional belief-based strategies, after a sufficiently long string of one-period histories " Cc ", both players think "the opponent is in the *cooperative* phase with very high probability". Now if, at this time, player h sees a one-period private history " Cd ", the two most likely events are, again, (1) signal " d " is an erroneous signal, (2) player l saw an erroneous signal " d " in the last period and is now in the *defective* phase. However, if we have $p_h \gg p_l$, player h now believes that most likely the signal " d " is an observation error since the probability that player l observes an error is very small compared to his. Alternatively, after a sufficiently long private history of the form " Cc, \dots, Cc ", a single bad signal " d " *does not decrease player h 's belief enough* as player h is much more likely to receive a wrong signal compared to player l . Hence, player h will have no incentive to punish player l after a bad signal " d ". **Lemma 1** formalizes the above arguments:

Lemma 1 *In the repeated prisoners' dilemma with asymmetric private monitoring, for any $p_h > 0$, we can always find a p_l and a q ($p_h > p_l > 0, q > 0$) such that the full belief-based strategy profile $((\rho_l, \hat{\pi}_h), (\rho_h, \hat{\pi}_l))$, defined in (2), is **not** a Nash equilibrium.*

To show the above result formally, we first define certain belief operators, which are the probabilities that the other player is in the *cooperative* state (in other words, the other player j has received a good private history, " Cc, \dots, Cc ", and will continue with σ_j^C), given last-period's belief μ . Denote $\chi_i^{Xy}(\mu)$ to be player i 's belief about the other player being in the cooperative

¹²Initial beliefs of the opponent (j)'s strategy being σ_j^C are $\hat{\pi}_l$ and $\hat{\pi}_h$. Also, χ_i^* is the fixed point of the belief operator $\chi_i^{Cc}(\mu)$ (defined in (3)).

¹³It should be noted that Sekiguchi (1997) also mentioned (*Remark 2* of the main theorem) that when a player's probability of receiving a wrong signal is greater than that of the opponent, the strategy profile he constructed *may* fail to be an equilibrium.

state, given that player i played X and saw signal y in this period and had belief μ in the previous period.

The evolution of player h 's beliefs on the equilibrium path is (partly) determined by the following belief operators (where player h 's initial belief is π_l):

$$\begin{cases} \chi_h^{Cc}(\mu) = \frac{(1-p_h-p_l-q)\mu}{(1-p_h-q)\mu+(p_h+q)(1-\mu)}; \\ \chi_h^{Cd}(\mu) = \frac{h\mu}{(p_h+q)\mu+(1-p_h-q)(1-\mu)}. \end{cases} \quad (3)$$

We further define criterion function

$$\Delta V_i(\mu_i; \delta, p_h, p_l, q) = \mu_i (V_i^{CC} - V_i^{DC}) - (1 - \mu_i) (V_i^{DD} - V_i^{CD}) \quad (4)$$

to be *the payoff difference* for player i between playing σ_i^C and playing σ_i^D , believing the opponent is in the cooperative state (having seen a good private history and continuing with σ_j^C) with probability μ_i . Note that $\Delta V_i(\mu_i; \delta, p_h, p_l, q)$ is linear and increasing in μ_i if $\delta > \frac{G}{1+G}$ and p_h, p_l and q are small. There is hence a *unique* vector $(\mu_l, \mu_h) = (\hat{\pi}_h, \hat{\pi}_l)$ such that $\Delta V_l(\mu_l; \delta, p_h, p_l, q) = \Delta V_h(\mu_h; \delta, p_h, p_l, q) = 0$, and " $\Delta V_i(\mu_i; \delta, p_h, p_l, q) > 0$ (< 0)" implies that "given i 's belief μ_i and j 's strategy, player i 's best response is σ_i^C (σ_i^D)". Therefore, the initial randomization probability $\hat{\pi}_j$ can also be interpreted as the *threshold* where player i 's behavior changes.

With the belief operators and the criterion function defined above, we now analyze player h 's beliefs and incentives after a private history of the form " $\underbrace{Cc, \dots, Cc, Cd}_n$ " where n is very large.

Denote player h 's belief of player l being cooperative (or player l has seen " Cc, \dots, Cc " and will continue with σ_l^C) after a private history " Cc, \dots, Cc, Cd " to be $\chi_h(Cc, \dots, Cc, Cd)$, which can be calculated iteratively from the belief operators defined in (3). Fixing $p_h > 0$, we can derive¹⁴ that as $p_l, q \rightarrow 0$, which can be interpreted as "the private monitoring structure becomes more and more asymmetric",

$$\chi_h \left(\underbrace{Cc, \dots, Cc, Cd}_n \right) \rightarrow 1, \text{ if } n \text{ is arbitrarily large.} \quad (5)$$

On the other hand, we can also show that there is some number $k > 0$ such that:

$$\lim_{p_l, q \rightarrow 0} \hat{\pi}_l = k < 1. \quad (6)$$

Intuitively, given any $p_h > 0$, player h is subject to observation errors in each period. Thus, to avoid the problem of not willing to punish a bad signal " d " at the end of the first period for player h (when private signals are conditionally independent), player l has to randomize with positive probability on σ_l^D at $t = 0$.

Results (5) and (6) indicate that when the " n " in player h 's private history " $\underbrace{Cc, \dots, Cc, Cd}_n$ " is large enough and p_l, q are sufficiently smaller than p_h , we will have that $\chi_h(Cc, \dots, Cc, Cd) > \hat{\pi}_l$, or after a long period of cooperation, player h does not want to punish a bad signal " d ", which is an informative indicator of l having defected. As the private history of the form " Cc, \dots, Cc, Cd "

¹⁴See Appendix for a more detailed argument.

happens with positive probability on the equilibrium path, the conventional belief-based strategy profile is *not* a Nash equilibrium when the private monitoring structure is asymmetric enough.

Although it seems that we need the condition " $p_h \gg p_l, q$ " to derive the result in **Lemma 1**, this condition is imposed for simplicity. We may obtain the above result under milder conditions, especially when players signals are negatively correlated. In addition, **Lemma 1** holds *no matter* how small p_h , p_l and q are, provided that p_l and q are sufficiently smaller than p_h . Thus, the result in **Lemma 1** is always relevant for the analysis of *almost-perfect* private monitoring structures.

2.3 Modified Belief-Based ("*Keep Them Guessing*") Strategies for Asymmetric Private Monitoring

Lemma 1 is not a particularly surprising result since after all, one should not expect a given set of strategies to continue to work when the information structure changes. However, this result *does* pose a natural and interesting question: are there other belief-based strategies that are "immune" to such asymmetric private monitoring imperfections? That is, are there sequential equilibria in some belief-based strategies (different from (2)) such that the efficient payoff profile (1, 1) can be approximated when the observation errors (p_h, p_l and q) are small enough? This section provides a positive answer to the above question. Specifically, we construct certain belief-based strategy (called "modified belief-based strategy" or "keep them guessing strategy" *interchangeably* hereafter) profiles where player l is required to randomize more than once on the equilibrium path so as to avoid the problem in **Lemma 1**.

It is helpful to first briefly describe the main idea of the construction. First, note that the pure "grim-trigger" profile is again not a Nash equilibrium under the asymmetric (conditionally independent) private monitoring setting: when a player sees a signal " d " in the first period, she believes that her opponent is still cooperating and thus does not want to initiate the punishment phase. Again, randomization in the initial period introduces small uncertainty (that the realization of the other's initial randomization might be " D ") for the players to resolve this "*unwilling-to-punish*" problem. In addition, **Lemma 1** shows that as the game proceeds, a *similar* problem arises: after a history of the form " Cc, \dots, Cc, Cd ", player h does not want to switch to the permanent defective phase. Motivated by the role of initial randomizations in resolving the problem of not willing to punish the first-period bad signals, a possible way to fix player h 's incentive problem is to let player l randomize between " C " and " D " when $t = 0$ and whenever player l has a history of the form " Cc, \dots, Cc ", so that player h can never be *too* sure about "player l being in the cooperating state" (keeping player h guessing). Now when player h sees a signal " d " after a long period of cooperation, he believes that the most likely events are "*he has received an erroneous signal*", "*player l played C in the last period and saw a signal d* " and "*player l 's last-period randomization has a realization D* ". The randomization probabilities and the discount factor can be properly chosen so that player l is indifferent after a history " Cc, \dots, Cc " and player h is willing to defect after a history " Cc, \dots, Cc, Cd ".

The above modification is, however, nontrivial: *first*, to evaluate equilibrium payoffs and to study the evolution of players' beliefs, a certain stationarity in the strategy of player l is important; *second*, player l 's incentive to randomize after a class of private histories should always be maintained, which imposes a restriction on the evolution of player l 's beliefs as we will see; *third*, the initial probabilities and the probability with which player l plays " C " in each period after a private history " Cc, \dots, Cc " must be carefully specified so that for any player, at any period, the incentives to punish bad signals are always sustained. In what follows, we

construct modified belief-based strategies that satisfy all these requirements, starting by defining partial continuation strategy sets for the two players.

2.3.1 The Partial Continuation Strategies for the Players

As in Bhaskar and Obara (2002), we first define and analyze the key components in constructing our modified belief-based strategies, *partial continuation strategies*. A partial continuation strategy for a player only specifies the player's behavior on the equilibrium path, but *not* off the equilibrium path. Partial continuation strategies provide building blocks for the construction of full belief-based strategies (to be defined) and they also play important roles in showing that the full belief-based strategy profile constitutes a sequential equilibrium.

Player h 's partial continuation strategies are " σ_h^C " and " σ_h^D ", while player l 's partial continuation strategies are defined as " σ_l^D " and some *mixed partial continuation strategy*: instead of playing " C " for sure after a history of the form " Cc, \dots, Cc ", player l randomizes between " C " and " D ", with some fixed probability p on " C ". Specifically, the set of partial continuation strategies for player h is $\{\sigma_h^C, \sigma_h^D\}$. For player l , we modify the set of partial continuation strategies to be:

$$\begin{aligned} \sigma_l &= \{\sigma_l^D, \sigma_l^M\}, \text{ where } (p \text{ to be determined}) \\ \sigma_l^M &: \begin{cases} \text{play } p \circ C + (1-p) \circ D, \text{ if } t = 0; \\ \text{play } \sigma_l^M \text{ after } "Cc", \text{ while play } \sigma_l^D \text{ after } "Cd, Dc \text{ or } Dd", \text{ if } t > 0. \end{cases} \end{aligned} \quad (7)$$

We can see that when the probability p equals 1 in σ_l^M , the partial continuation strategy σ_l^M is identical to σ_l^C .

Now, as described before, to resolve the problem in **Lemma 1**, we can let player l randomize all the time between C and D along a particular sequence of histories so as to keep introducing some uncertainty to player h . To construct such a strategy for player l , it is convenient to let player l always randomize between C and D in the same way over time in σ_l^M . That is, player l plays C with the *same* probability p whenever l sees a private history of the form " Cc, \dots, Cc ", or p is time-independent. If such a mixed strategy can be arranged for l , the dynamics of player h 's beliefs in the equilibrium will be much more tractable. Moreover, such a construction will simplify the calculations of players' *ex ante* equilibrium payoffs. On the other hand, such an arrangement also brings some tension, since first-period randomization is determined by δ and (p_h, p_l, q) , while given h 's strategy, whenever player l sees a one-period private history " Cc ", her belief (that h 's continuation strategy is " σ^C ") increases and it thus seems impossible for player l to be always indifferent between C and D after such a history. One can, for each (p_h, p_l, q) , fix δ^{15} , so that player l 's belief about h 's continuation strategy after each " Cc " is always at its fixed point, thus always invariant after each additional one-period history " Cc ". This way, player l can possibly randomize all the time with the same probability " p " on C . In addition, the probability " p " has to be specified properly so that player h 's incentives are correct in any information set h may have.

In summary, the equilibrium path is designed as follows: player h randomizes between σ_h^C and σ_h^D and player l randomizes between σ_l^M and σ_l^D in the initial period. After that, player h plays σ_h^C whenever his private history is of the form " Cc, Cc, \dots, Cc ", and plays σ_h^D otherwise. Player l plays σ_l^M whenever she sees " Cc, Cc, \dots, Cc ". If her randomization realizes D or if she ever sees a signal " d " on the equilibrium path, she will play σ_l^D for the rest of the game. Player

¹⁵We use the public randomization device to extend the result to cases where players are more patient.

l 's randomization after the initial period always has the form " $p \circ C + (1 - p) \circ D$ " (in σ_l^M) so as to simplify the analysis on the belief dynamics and calculations of equilibrium payoffs.

As we will see shortly, our modified strategy is still a *belief-based strategy* since for each player, at any stage of the game, the support of the player's beliefs always contains the same partial continuation strategies: " σ_h^C " and " σ_h^D " (or " σ_l^M " and " σ_l^D "). And, as we will see, given any belief of player i ($i \in \{h, l\}$), player i 's best response is also either " σ_h^C / σ_l^M " or " σ_h^D ". Moreover, a player's belief of the continuation strategy of the opponent is a *sufficient statistic* for the player's private histories and all inferences about the opponent's future behavior can thus be based on the player's beliefs (hence we have to study the dynamics of the players' beliefs to check for sequential rationality as in Sekiguchi (1997) and Bhaskar and Obara (2002)).

In what follows, we outline the main steps of constructing the modified belief-based strategies, leaving the details to Appendix. First, as argued above, initial randomizations are still necessary. The initial randomization probabilities (π_l, π_h) can be calculated similarly using the following "coordination game" form as in **Figure 2**:

	σ_h^C	σ_h^D
σ_l^M	V_l^{MC}, V_h^{CM}	V_l^{MD}, V_h^{DM}
σ_l^D	V_l^{DC}, V_h^{CD}	$0, 0$

FIGURE 3. The "coordination game" at $t = 0$.

Next, we derive certain conditions on δ and (p_h, p_l, q) so that (1) player l is willing to randomize as required after a private history of the form " Cc, \dots, Cc ", and (2) the initial randomization probabilities (π_l, π_h) and the probability p are well-defined ($0 < \pi_l, \pi_h, p < 1$). We start by analyzing the situation for player l :

(A) Given player h 's strategy, player l 's belief after each private one-period history " Cc " is:

$$\chi_l^{Cc}(\mu) = \frac{(1 - p_h - p_l - q)\mu}{(1 - p_l - q)\mu + (p_l + q)(1 - \mu)},$$

where μ is player l 's belief of player h 's continuation strategy being " σ_h^C " at the end of last period. The belief operator $\chi_l^{Cc}(\mu)$ has an *interior* fixed point $\tilde{\chi}_l^{Cc}$:

$$\tilde{\chi}_l^{Cc} = \frac{1 - p_h - 2p_l - 2q}{1 - 2p_l - 2q}. \quad (8)$$

Now, for player l to be willing to randomize both in the initial period and after seeing " Cc, \dots, Cc ", one possibility is that player l 's *initial* belief (that player h will play σ_h^C), π_h , be the same as her belief after a private history of the form " Cc, \dots, Cc ". We thus impose the restriction that the probability π_h in player h 's initial randomization be *identical* to **the fixed point** ($\tilde{\chi}_l^{Cc}$) of the belief operator $\chi_l^{Cc}(\mu)$ of player l . In particular, this implies that player l 's belief of h 's *private state* does not change over time whenever player l has always played " C " and has always seen signal " c " previously. We thus have our first restriction as:

$$\chi_l^{Cc}(\pi_h) = \frac{1 - p_h - 2p_l - 2q}{1 - 2p_l - 2q},$$

which reduces to

$$\pi_h = \tilde{\chi}_l^{Cc} = \frac{1 - p_h - 2p_l - 2q}{1 - 2p_l - 2q}. \quad (9)$$

(B) For the strategy profile to be well-defined, the probabilities π_l and p have to lie between 0 and 1. It will be clear that certain conditions have to be imposed on (p_h, p_l, q) so that $\pi_l, p \in (0, 1)$ ¹⁶. In addition, since compared to the conventional belief-based strategies ($p = 1$), constructing any $p \in (0, 1)$ in the strategy introduces certain efficiency loss, the probability p should converge to 1 as (p_h, p_l, q) goes to $\mathbf{0}$ so that efficiency loss is minimal and full efficiency can be approximated when observation errors are *arbitrarily* small. With these two requirements in mind, we impose the following restriction : "the probability in player l 's initial randomization is equal to the probability p in σ_l^M ", or,

$$\pi_l = p. \quad (10)$$

Recall that π_l is the initial probability that player l plays σ_l^M , while p is the probability of playing C in σ_l^M for player l . Although p and π_l are related, condition (10) is not a necessary condition for our results and there are other choices of p for our construction to work. However, condition (10) will simplify our analysis later and as long as we can show that p (hence π_l) is between 0 and 1, we only need to keep track of a same probability π_l afterwards.

Lemma 2 shows that the set of restrictions (9) and (10) implicitly defines a *unique* $p \in (0, 1)$ and a *unique* $\delta^* \in (0, 1)$ as functions of (p_h, p_l, q) , provided that p_h, p_l , and q are sufficiently small. The purpose of **Lemma 2** is to show that the two restrictions imposed above do not yield conflicting results for p and δ when observation errors are small enough.

Lemma 2 *The set of equations (9) and (10) implies that:*

1. *There exists a neighborhood, \mathbf{U} , of (p_h, p_l, q) near the origin such that δ^* is a unique C^1 function¹⁷ of (p_h, p_l, q) in \mathbf{U} . Moreover, $\delta^* \in (0, 1)$ and $\frac{\partial \delta^*}{\partial p_h}, \frac{\partial \delta^*}{\partial p_l}$ and $\frac{\partial \delta^*}{\partial q} > 0$ for all $(p_h, p_l, q) > \mathbf{0}$ in \mathbf{U} with $p_h > p_l$.*

2. *There exists a neighborhood, \mathbf{V} , of (p_h, p_l, q) near the origin such that p is a unique C^1 function of (p_h, p_l, q) in \mathbf{V} . Moreover, $p \in (0, 1)$ and $\frac{\partial p}{\partial p_h}, \frac{\partial p}{\partial q} < 0$ for all (p_h, p_l, q) in \mathbf{V} with $p_h > p_l$.*

Proof. See Appendix. ■

In addition, note that in the limit, we obtain

$$\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} \delta^* = \frac{G}{1+G} \quad \text{and} \quad \lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} \pi_l = \lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} \pi_h = 1. \quad (11)$$

We have two remarks for conditions (9) and (10) :

1. By **Lemma 2**, it is now clear that as long as p_h, p_l and q are small enough, δ^* is well-defined and there exists a $p \in (0, 1)$ for us to construct the strategy profile. On the other hand, the restriction of p_h, p_l , and q being small does not pose a problem for us since our main concern in this paper is again to investigate the repeated prisoners' dilemma in asymmetric, but *almost-perfect* private monitoring settings.

¹⁶Note that the role of p in the strategy profile is to provide sufficient incentives for player h to defect when he sees a bad signal. As we will see in **Lemma 4**, condition " $p \in (0, 1)$ " is the only requirement on p to construct the sequential equilibrium. However, to demonstrate that full efficiency (1,1) can be approximated using the modified belief-based strategies for our second main result, we will also need condition " $\lim_{(h,l,q) \rightarrow \mathbf{0}} p = 1$ ".

¹⁷Define this function to be $\delta^*(h, l, q)$.

2. Condition (9) ensures that after a good history " Cc, \dots, Cc ", player l 's belief of player h being in the cooperative state is identical to her belief in the initial period. Given this condition, player l does find it optimal to randomize, both in the initial period and after a private history of the form " Cc, \dots, Cc ". However, one might think that condition " $\pi_h = \tilde{\chi}_l^{Cc}$ " makes player l 's strategy a bit unreasonable since player l is now required to randomize *differently* when she is in a same information set. That is, in the initial period, where she believes that player h 's continuation strategy is " $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$ ", she plays " $\pi_l \circ \sigma_l^M + (1 - \pi_l) \circ \sigma_l^D$ ", while after the initial period and a private history of the form " Cc, \dots, Cc ", player l plays " σ_l^M ", when again, her belief of player h 's continuation strategy is " $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$ " (by condition (9)). This is related to the purifiability issue of the modified belief-based strategies, and will be discussed in more detail at the end of **Section 2.4**.

2.3.2 Evolution of the Beliefs of the Players

Since players face different observation errors and adopt different partial continuation strategies, players' beliefs now evolve in a more complicated way than that in Bhaskar and Obara (2002).

Specifically, adopting the notation from Bhaskar and Obara (2002), we define certain belief operators for the players, which are again the probabilities that the other player is in the cooperative state¹⁸ given last-period's belief μ . As before, denote $\chi_i^{Xy}(\mu)$ to be player i 's belief given that i played X and saw signal y in this period and had belief μ in the previous period. Note that to calculate player h 's beliefs on player l being in the cooperative state (player l 's continuation strategy is σ_l^M), player l 's realization of last-period randomization must be " C ".

Using Bayes' rule, we have:

- Player h :

$$\begin{aligned}
\chi_h^{Cc}(\mu) &= \frac{(1 - p_h - p_l - q) \mu \pi_l}{(1 - p_h - q) \mu \pi_l + (p_h + q) (1 - \mu \pi_l)}; \\
\chi_h^{Cd}(\mu) &= \frac{p_h \mu \pi_l}{(p_h + q) \mu \pi_l + (1 - p_h - q) (1 - \mu \pi_l)}; \\
\chi_h^{Dc}(\mu) &= \frac{p_l \mu \pi_l}{(1 - p_h - q) \mu \pi_l + (p_h + q) (1 - \mu \pi_l)}; \\
\chi_h^{Dd}(\mu) &= \frac{q \mu \pi_l}{(p_h + q) \mu \pi_l + (1 - p_h - q) (1 - \mu \pi_l)}.
\end{aligned} \tag{12}$$

- Player l :

$$\begin{aligned}
\chi_l^{Cc}(\mu) &= \pi_h = \frac{1 - p_h - 2p_l - 2q}{1 - 2p_l - 2q}; \\
\chi_l^{Cd}(\mu) &= \frac{p_l \mu}{(p_l + q) \mu + (1 - p_l - q) (1 - \mu)}; \\
\chi_l^{Dc}(\mu) &= \frac{p_l \mu}{(1 - p_l - q) \mu + (p_l + q) (1 - \mu)}; \\
\chi_l^{Dd}(\mu) &= \frac{q \mu}{(p_l + q) \mu + (1 - p_l - q) (1 - \mu)}.
\end{aligned} \tag{13}$$

¹⁸Note that "being in the cooperative state" now means differently for different players. For player l , $\chi_l^{Xy}(\mu)$ denotes l 's subjective probability that player h will play σ_h^C after one-period private history Xy , while for player h , $\chi_h^{Xy}(\mu)$ denotes h 's subjective probability that player l will play σ_l^M after Xy .

Note that the initial belief vector $\boldsymbol{\mu}$ is (π_h, π_l) (recall that the initial randomization probability vector is (π_l, π_h)). We now present **Lemma 3**, which summarizes certain properties of the belief operators defined in (12) and (13).

Lemma 3 *The following is true for the belief operators defined in (12) and (13) :*

1. The belief operator $\chi_h^{Cc}(\boldsymbol{\mu})$ has a fixed point defined by

$$\chi_h^* = \frac{(1 - p_h - p_l - q) \pi_l - p_h - q}{(1 - 2p_h - 2q) \pi_l}. \quad (14)$$

In addition, in the open neighborhood $\mathbf{U} \cap \mathbf{V}$ near $(0, 0, 0)$, we can choose some $p = \pi_l < 1$ by **Lemma 2**, and thus $\chi_h^* < 1$.

2. There exists some open set $\mathbf{W} \subset \mathbb{R}^3$ near $(0, 0, 0)$ such that $\pi_l < \chi_h^*$ for all $(p_h, p_l, q) \in \mathbf{W}$.
3. $\frac{d\chi_h^{Cc}(\boldsymbol{\mu})}{d\boldsymbol{\mu}} > 0$, $\frac{d\chi_h^{Cd}(\boldsymbol{\mu})}{d\boldsymbol{\mu}} > 0$, and $\frac{d\chi_l^{Dc}(\boldsymbol{\mu})}{d\boldsymbol{\mu}}, \frac{d\chi_l^{Dd}(\boldsymbol{\mu})}{d\boldsymbol{\mu}} > 0$.

Proof. See Appendix. ■

Lemma 3 is prepared for our next important result (**Lemma 4**): "*the modified belief-based strategies (to be defined shortly) are realization equivalent to the partial continuation strategies σ defined previously*", which is, in turn, essential for proving our first main result that the modified belief-based strategy profile is a sequential equilibrium in the following section.

2.3.3 The Full Modified Belief-Based Strategy Profile

Now we formally define the players' *full* modified belief-based strategies for the asymmetric private monitoring setting.

Recall that $\sigma_h = \{\sigma_h^C, \sigma_h^D\}$ and $\sigma_l = \{\sigma_l^M, \sigma_l^D\}$ are the sets of partial continuation strategies for player h and player l , respectively. On the equilibrium path, players play the following:

- Player h :

"Play $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$ in the initial period. Starting from the second period, player h continues with the strategy σ_h^C / σ_h^D that was realized at $t = 0$ ".

- Player l :

"Play $\pi_l \circ \sigma_l^M + (1 - \pi_l) \circ \sigma_l^D$ in the initial period. Starting from the second period, player l continues with the strategy σ_l^M / σ_l^D that was realized at $t = 0$ ".

Now the *full* modified belief-based strategies are defined as follows: *player i plays according to the pair (ρ_i, π_j) , where $(\chi_h^*$ is defined in (14))*

$$\left\{ \begin{array}{l} \rho_h: [0, \chi_h^*] \rightarrow \{C, D, \pi_h \circ C + (1 - \pi_h) \circ D\} \\ \text{s.t. } \rho_h(\boldsymbol{\mu}) = \begin{cases} C, & \text{if } \mu > \pi_l; \\ D, & \text{if } \mu < \pi_l; \\ \pi_h \circ C + (1 - \pi_h) \circ D, & \text{if } \mu = \pi_l. \end{cases} \end{array} \right. \quad (15)$$

$$\left\{ \begin{array}{l} \rho_l: [0, \pi_h] \rightarrow \{C, D, \pi_l \circ C + (1 - \pi_l) \circ D, \pi_l^2 \circ C + (1 - \pi_l^2) \circ D\} \\ \text{s.t. } \rho_l(\boldsymbol{\mu}) = \begin{cases} D, & \text{if } \mu < \pi_h; \\ \pi_l^2 \circ C + (1 - \pi_l^2) \circ D, & \text{if } \mu = \pi_h \text{ and } t = 0; \\ \pi_l \circ C + (1 - \pi_l) \circ D, & \text{if } \mu = \pi_h \text{ and } t \geq 1. \end{cases} \end{array} \right. \quad (16)$$

It is worth noting that for player h , given any $\mu_h \in [0, \chi_h^*]$, the belief operators " $\chi_h^{Cc}(\mu)$, $\chi_h^{Cd}(\mu)$, $\chi_h^{Dc}(\mu)$, $\chi_h^{Dd}(\mu)$ " have ranges as some subsets of $[0, \chi_h^*]$. For the belief operator $\chi_h^{Cc}(\mu)$, this is true if $\pi_l < 1$. For the other belief operators, note that all of the curves " $\chi_h^{Cd}(\mu)$, $\chi_h^{Dc}(\mu)$, $\chi_h^{Dd}(\mu)$ " lie below the 45-degree line and hence lie below $\chi_h^{Cc}(\mu)$. It is therefore sufficient for us to only consider beliefs in $[0, \chi_h^*]$ for player h at *any* time. Similarly, for player l , we only need to consider l 's beliefs in $[0, \pi_h]$.

Our next step is to show that the belief-based strategy (ρ_i, π_j) , $i \in \{h, l\}$, is *realization equivalent*¹⁹ to the partial continuation strategies defined in (7). This is presented in **Lemma 4**.

Lemma 4 *Suppose that $(p_h, p_l, q) \in \mathbf{W}$ ²⁰. The modified belief-based strategies (ρ_l, π_h) and (ρ_h, π_l) are realization equivalent to σ_l and σ_h , respectively.*

Proof. See Appendix. ■

Now we are ready to apply **Lemma 4** to show our first main result: *the modified belief-based strategy profile defined in (15) and (16) is a sequential equilibrium of the infinitely repeated prisoners' dilemma with asymmetric private monitoring, given that (p_h, p_l, q) are chosen from a small open neighborhood near the origin.* As in Bhaskar and Obara (2002), the proof of the result mainly employs some criterion functions (defined in Appendix) to show that there is no profitable one-shot deviation for player i at any information set if player i uses the specified partial continuation strategy σ_i , which is realization equivalent to (ρ_i, π_j) by **Lemma 4**.

Proposition 1 *There exists some open set $\mathbf{Z} \subset \mathbb{R}^3$ near $(0, 0, 0)$ such that for any given $(p_h, p_l, q) \in \mathbf{Z}$, the modified belief-based strategy profile $((\rho_l, \pi_h), (\rho_h, \pi_l))$ is a sequential equilibrium of the repeated prisoners' dilemma with asymmetric private monitoring.*

Proof. See Appendix. ■

We have four remarks for **Proposition 1**:

First, the modified belief-based (*keep them guessing*) equilibrium strategies in **Proposition 1** have certain interpretations on players' equilibrium behavior under asymmetric private monitoring. Specifically, the equilibrium strategies indicate that to sustain cooperation in an asymmetric private monitoring setting, the better-informed player (player l) has to act more erratically by defecting sometimes so that the less-informed player's (player h) private signals are *always sufficiently informative* about the better-informed player's actions and thus the less-informed player would have enough incentives to punish the opponent whenever he sees a bad signal. A comparison of this equilibrium behavior with observed behavior in real applications would be interesting to investigate in future work.

Second, the entire modified belief-based strategy profile was constructed on the condition of (p_h, p_l, q) being close enough to the origin so that δ^* and π_l (thus p) are well-defined and the beliefs of player h evolve appropriately (player l 's beliefs evolve in a similar way as that in Bhaskar and Obara (2002)). This condition is, however, not a strong one and has been imposed in most of the literature on repeated games with private monitoring, the primary concern of

¹⁹Two strategies of a player i , σ_i and σ'_i are *realization equivalent*, if (σ_i, σ_{-i}) and (σ'_i, σ_{-i}) induce the same probability distribution over the outcomes, other players' strategies being fixed and given by σ_{-i} .

²⁰Recall that \mathbf{W} is defined in **Lemma 3**.

which (so far) being whether it is possible to construct sequential equilibria to support non-myopic behavior when the history of play is *almost common knowledge*²¹.

Third, in **Proposition 1**, the belief-based strategy profile $((\rho_l, \pi_h), (\rho_h, \pi_l))$ is a sequential equilibrium *only* when players have a fixed discount factor δ^* defined in **Lemma 2**, which is derived from the restriction $\pi_h = \tilde{\chi}_l^{Cc}$. As discussed before, this restriction is imposed so that player l 's belief is fixed after any history of the form " Cc, \dots, Cc ", and player l is therefore willing to randomize both in the first period and after a private history of the form " Cc, \dots, Cc ". To extend **Proposition 1** to the situation where the players are more patient ($\delta > \delta^*$), we can²² follow Bhaskar and Obara (2002) to use a public randomization device to reduce the common discount factor δ *exactly* to δ^* for any $\delta > \delta^*$ (see **Lemma 5** in the next section).

Four, the modified belief-based strategy profile in **Proposition 1** requires player l to play differently at the same information set, which might be considered somewhat unreasonable. That is, player l has the same belief of h 's continuation strategy " $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$ " at $t = 0$ and after l 's private history of the form " Cc, \dots, Cc ". But player l is required to play " $\pi_l \circ \sigma_l^M + (1 - \pi_l) \circ \sigma_l^D$ " at $t = 0$ and to play " σ_l^M " after " Cc, \dots, Cc ". There is no incentive problem for l , since in both cases, player l is indifferent between σ_l^M and σ_l^D . But the randomizations constructed above may raise some suspicion that the equilibrium may not be purifiable, or the equilibrium may not be robust to payoff perturbations (Harsanyi (1973)). Mixed-strategy equilibria for generic normal-form games can be justified by Harsanyi's purification theorem where Harsanyi (1973) shows that every mixed-strategy equilibrium is the limit of pure-strategy equilibria in incomplete-information games with payoff perturbations. However, little work in the literature has studied purification for extensive-form games. Bhaskar (1998) shows in an overlapping generation game that a mixed strategy equilibrium cannot be purified when payoff shocks are in a time separable form. Bhaskar (1998) argues that the non-purifiability comes from the fact that players randomize differently in the same information set, or the randomizing probabilities depend on payoff-irrelevant histories. This is also a potential criticism of the belief-free strategies in which a lot of randomizations are involved in the construction and players are required to randomize differently at the same information set. Bhaskar, Mailath and Morris (2006) study the purifiability of the class of one-period memory mixed strategy equilibria used by Ely and Välimäki (2002), where they show that none of such mixed-strategy equilibria are the limit of one-period memory equilibria of the perturbed games, for almost all noise distributions. However, for mixed-strategy equilibria of repeated games where the strategies have infinite memory (in other words, have infinite history dependence), there is so far *no* evidence in the literature to show such mixed-strategy equilibria are not purifiable when players randomize differently in the same information set. Our modified belief-based strategies are mixed strategies that have infinite history dependence as players' beliefs are crucial components of the strategies and the beliefs are calculated *recursively* from all previous private histories. Hence, according to the literature, it is not clear whether our modified strategies are purifiable or not. We come back to this issue in **Section 4** where we make a detailed comparison between belief-free equilibria

²¹A recent paper by Matsushima (2004) is, however, an exception. Matsushima (2004) uses a "review phase" (introduced by Radner (1985)) and extends the two-state strategies of Ely and Välimäki (2002) to establish a Folk theorem for the prisoners' dilemma with conditionally independent private monitoring, but without vanishing observation errors.

²²Another possibility is to use "Ellison's trick" to divide the original repeated game into n ($n \in \mathbb{N}$) distinct repeated games to obtain a lower *effective* discount factor for the players. However, it is almost always not possible to obtain the discount factor δ^* from $\delta > \delta^*$ because of the "integer" problem (it is rarely the case that $\delta^n = \delta^*$ for $\delta > \delta^*$).

and belief-based equilibria.

2.4 Approximating the Efficient Payoff Pair (1, 1) when Observation Errors are Arbitrarily Small

As stated in the *second* remark at the end of the previous section, a public randomization device is necessary for us to extend the results in **Proposition 1** to the situation of $\delta > \delta^*$. **Lemma 5**, borrowed from Bhaskar and Obara (2002), shows that we can always construct a belief-based sequential equilibrium, as in **Proposition 1**, for the repeated game with asymmetric private monitoring when players have a common discount factor $\delta > \delta^*$, if a public randomization device uniformly distributed on $[0, 1]$ is available for the players.

Lemma 5 (*Lemma 3 in Bhaskar and Obara (2002)*) *Suppose that there is a strategy profile which is a sequential equilibrium, yielding payoffs (v_1, v_2) for some $\delta^* \in (0, 1)$. If a public randomization device uniformly distributed on the unit interval is available for the players, then (v_1, v_2) is a sequential equilibrium payoff profile for any $\delta > \delta^*$.*

The equilibrium payoffs for the players when p_h, p_l and q are arbitrarily small can be calculated from expressions V_h^{CM} and V_l^{MC} (calculated in **Section 5.2** in Appendix):

$$\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} V_h^{CM} = 1 \text{ and } \lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} V_l^{MC} = 1. \quad (17)$$

These are true by the fact that $\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} \delta^* = \frac{G}{1+G}$ and $\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} \pi_l = \lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} \pi_h = 1$ (expression (11)).

We now show that the symmetric payoff pair (1, 1) can be approximated when the observation errors are *arbitrarily* small. This is presented in **Proposition 2**:

Proposition 2 *For the repeated prisoners' dilemma with asymmetric private monitoring, given any $v < 1$, there is a sequential equilibrium in belief-based strategies with payoffs (v_l, v_h) where $v_l, v_h \in (v, 1)$, provided that (1) (p_h, p_l, q) is close enough to $(0, 0, 0)$ and (2) there is a public randomization device and players have a common discount factor $\delta \geq \delta^*$.*

Proof. See Appendix. ■

3 Extensions of the "Keep Them Guessing" Strategies to General Games - An Example

The "keep them guessing" strategies constructed in the previous sections can be used to construct non-trivial equilibria for other games. In this section, we extend the method developed in the previous part to show that our construction has the potential to be generalized to other two-player repeated games, with certain restrictions on payoffs of the stage game. We do this by constructing a cooperative sequential equilibrium (not repetitions of stage game Nash equilibria) of a specific symmetric two-player repeated game example.

	C	D	E
C	1, 1	-1, 0	-2, 3
D	0, -1	0, 0	-1, -1
E	3, -2	-1, -1	-2, -2

FIGURE 4. Original example in Sekiguchi (1997).

The main idea of the example (to be constructed) is motivated by the following observation in Sekiguchi (1997): consider now the repeated game, whose stage game is defined in **Figure 4**, under *symmetric private monitoring*. Sekiguchi used the example in **Figure 4** to show that it is difficult to generalize the mixed strategy profile (mixture of " σ^C " and " σ^D ") he constructed to other games. The argument is that if a player sees a private history " Cc, \dots, Cc, Ce ", the player will regard the bad signal " e " as an observation error (if the private monitoring structure is, for example, conditionally independent) and will therefore not want to switch to " σ^D " forever. This destroys the incentive to cooperate for the other player. This problem of unwilling to punish bad signals can be similarly solved if we again construct some ongoing randomizations for the players along the cooperative path so as to keep introducing small amounts of uncertainty for the players. The ongoing randomization probability can be chosen properly so that (1) players can be induced to punish the opponent after a bad signal (here " d " or " e ") and (2) efficiency loss from the ongoing randomizations is small. In other words, we can similarly construct a "mixture" trigger (partial continuation) strategy, σ^M , instead of σ^C , in the first-period randomization. Now a bad signal " e " can either mean an observation error or the realization of the other player's last-period randomization being " E ". That way, players would have incentives to punish bad signals if they are not very patient (same for the bad signal " d ").

Specifically, the main idea of the potential extension is the following: for general two-player repeated games with full-support (almost-perfect) private monitoring, we can employ a set of partial continuation strategies $\sigma = \{\sigma^M, \sigma^D\}$ to construct nontrivial equilibria, where the mixture in σ^M would involve (possibly) all actions in the stage game, while σ^D is defined as "*always play a stage Nash equilibrium*". Specifically, players play " $\pi \circ \sigma^M + (1 - \pi) \circ \sigma^D$ " initially. After $t = 0$, each player plays σ^M only after *good* private histories (here " Cc, \dots, Cc "), and plays σ^D otherwise. The advantage of this construction is its **simplicity**: there are only two states (for any player) to analyze at any time of the game: the *cooperative* (σ^M) state and the *defective* state (σ^D). This will, in particular, simplify the analysis of players' belief dynamics. But as we shall see shortly, this construction also imposes certain restrictions on stage game payoffs and the private monitoring structure of the repeated game.

Consider a two-player repeated game whose stage game is defined in **Figure 5**.

	C	D	E
C	1, 1	$-1, \frac{11}{4}$	-2, 3
D	$\frac{11}{4}, -1$	0, 0	-1, -1
E	3, -2	-1, -1	-2, -2

FIGURE 5. Stage Game of the Example

One can see that this game is a modified version of the example in Sekiguchi (1997). Since the construction of using " σ^M " also imposes certain restrictions on the payoffs of the stage game, the payoffs from action profile " DC " have been changed from $(0, -1)$ in **Figure 4** to $(\frac{11}{4}, -1)$ in this example. Note that (D, D) is the *unique* stage game Nash equilibrium where players obtain $(0, 0)$.

Consider the following symmetric, *conditionally independent* private signal distribution where each player's set of private signals is $Y = \{c, d, e\}$:

$$\left\{ \begin{array}{l} \Pr(y_j = c | a_i = C) = 1 - 2\varepsilon \\ \Pr(y_j = d | a_i = C) = \varepsilon \\ \Pr(y_j = e | a_i = C) = \varepsilon \end{array} \right. \quad \left\{ \begin{array}{l} \Pr(y_j = d | a_i = D) = 1 - 2\varepsilon \\ \Pr(y_j = c | a_i = D) = \varepsilon \\ \Pr(y_j = e | a_i = D) = \varepsilon \end{array} \right. \quad \left\{ \begin{array}{l} \Pr(y_j = e | a_i = E) = 1 - 2\varepsilon \\ \Pr(y_j = c | a_i = E) = \varepsilon \\ \Pr(y_j = d | a_i = E) = \varepsilon \end{array} \right.$$

FIGURE 6. Symmetric Private Monitoring Distribution of the Players.

We now follow the main idea of the extension to construct a sequential equilibrium where players obtain approximately $(\frac{16}{47}, \frac{16}{47})$ under the private monitoring structure defined in **Figure 6**.

As before, we first define the set of partial continuation strategies for the players: $\sigma_i = \{\sigma_i^M, \sigma_i^D\}$, where (p and q to be determined)

$$\sigma^M : \begin{cases} \text{play } p \circ C + q \circ D + (1 - p - q) \circ E, & \text{if } t = 0; \\ \text{play } \sigma^M & \text{after "Cc",} \\ \text{while } \sigma^D & \text{after "Cd, Ce, Dc, Dd, De, Ec, Ed or Ee", if } t > 0. \end{cases} \quad (18)$$

Now we analyze the first-period randomizations for the players. At $t = 0$, players face the following "coordination" game:

	σ_2^M	σ_2^D
σ_1^M	V^{MM}, V^{MM}	V^{MD}, V^{DM}
σ_1^D	V^{DM}, V^{MD}	V^{DD}, V^{DD}

The "coordination game" at $t = 0$.

FIGURE 7.

As before, each player plays " $\pi \circ \sigma^M + (1 - \pi) \circ \sigma^D$ " at $t = 0$ (where $\pi = \frac{-V^{MD}}{V^{MM} - V^{MD} - V^{DM}}$). After the initial randomization, each player plays " σ^M " only after the "good" private history " Cc, \dots, Cc ", while plays " σ^D " otherwise.

Next, we check the incentives of the players. The main difficulty here is how to sustain players' incentives to randomize among C , D and E in every period after a good history " Cc, \dots, Cc ". Checking players' incentives in other private histories is analogous to the case of the repeated prisoners' dilemma. For the repeated game (in **Figure 5**) with the private monitoring structure defined in **Figure 6**, there are several non-trivial restrictions we have to impose so that the players will find it optimal to randomize as is required.

First, a player's belief of the other being in the cooperative state should be invariant whenever the player sees a good history " Cc, \dots, Cc ". This is again resolved by setting the probability on σ^M of the initial randomization equal to the fixed point of the belief operator $\chi^{Cc}(\mu)$.

$$\pi = \chi_{Cc}^* \tag{R-1}$$

where $\chi_{Cc}^* = \frac{(1-2\varepsilon)^2 p - \varepsilon}{(1-3\varepsilon)p}$ is the fixed point of the belief operator²³

$$\begin{aligned} \chi^{Cc}(\mu) &= \frac{\Pr(cc|CC) p \mu}{\left\{ \begin{array}{l} [\Pr(cc|CC) + \Pr(cd, e|CC)] p \mu + [\Pr(cc|CD) + \Pr(cd, e|CD)] (1 - \mu) \\ + [\Pr(cc|CD) + \Pr(cd, e|CD)] q \mu + [\Pr(cc|CE) + \Pr(cd, e|CE)] (1 - p - q) \mu \end{array} \right\}} \\ &= \frac{(1 - 2\varepsilon)^2 p \mu}{(1 - 2\varepsilon) \mu p + \varepsilon (1 - \mu p)}. \end{aligned}$$

Second, given that a player has seen a good history " Cc, \dots, Cc " (or equivalently, given that a player's belief of the opponent being cooperative is π), there should be no profitable one-shot

²³To save on notation, I denote, for example, " $\Pr(cd|CC) + \Pr(ce|CC)$ " simply as " $\Pr(cd, e|CC)$ ".

deviation for the player. We first calculate *the continuation values* of playing C , D and E after a good history. Denote these values to be V_C , V_D and V_E , respectively²⁴. For example, V_D is defined as the value to a player after seeing a good private history " Cc, Cc, \dots, Cc ", if she plays D in this period and then conform to the equilibrium strategy afterwards.

There is no profitable one-shot deviation after a good private history if

$$V_C = V_D = V_E. \quad (R-2)$$

Note that in the repeated prisoners' dilemma with asymmetric private monitoring in **Section 2**, each player only has two actions in each stage and *only* player l is required to do the ongoing randomizations, a similar restriction as $(R-2)$ does not prevent us from approximating full efficiency when observation errors are arbitrarily small ($\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} p = 1$). In the above example, each player has three possible actions in each stage and both players are playing the "*keep them guessing*" strategy, the "no profitable one-shot deviation" condition will restrict us from choosing probability p freely. As a result, full efficiency is *not feasible* even when ε is arbitrarily small (see **Proposition 3**).

After imposing these two constraints, it is then possible to apply the implicit function theorem to show that when ε is small, we can find a " p " near $\frac{4}{5}$, a " q " near $\frac{1}{5}$ and a " δ " near $\frac{10}{11}$, all continuously differentiable functions of ε in a neighborhood $(0, \bar{\varepsilon})$. Now the full belief-based strategy for each player is defined as:

$$\left\{ \begin{array}{l} \rho : [0, \pi] \rightarrow \left\{ \begin{array}{l} C, D, \pi p \circ C + (\pi q + 1 - \pi) \circ D + \pi(1 - p - q) \circ E, \\ p \circ C + q \circ D + (1 - p - q) \circ E. \end{array} \right\} \\ s.t. \rho(\mu) = \left\{ \begin{array}{l} D, \text{ if } \mu < \pi; \\ p \circ C + (\pi q + 1 - \pi) \circ D + \pi(1 - p - q) \circ E, \text{ if } \mu = \pi \text{ and } t = 0; \\ p \circ C + q \circ D + (1 - p - q) \circ E, \text{ if } \mu = \pi \text{ and } t \geq 1. \end{array} \right. \end{array} \right. \quad (19)$$

The belief-based strategy (ρ, π) is realization equivalent to the partial continuation strategy defined in (18). We can in addition show that the full belief-based strategy profile $((\rho, \pi), (\rho, \pi))$ constitutes a sequential equilibrium. In this sequential equilibrium, players obtain approximate equilibrium payoffs $(\frac{16}{47}, \frac{16}{47})$. This is presented in **Proposition 3**. We leave all other detailed arguments to Appendix.

Proposition 3 *Consider the repeated game with private monitoring as defined in Figure 5 and Figure 6. There is an $\bar{\varepsilon}$ s.t. for any $\varepsilon \in (0, \bar{\varepsilon})$, the belief-based strategy profile $((\rho, \pi), (\rho, \pi))$ is a sequential equilibrium of the repeated game. In addition, as $\varepsilon \rightarrow 0$, players' payoffs approach $(\frac{16}{47}, \frac{16}{47})$.*

We conjecture that our construction in the above example can be generalized to some class of two-player repeated games with almost-perfect private monitoring. The advantage of the above construction is its simplicity: in any period of the game, each player can only be in one

²⁴Sepcifically, the continuation values V_C , V_D and V_E can be calculated as:

$$\begin{aligned} V_C &= \pi p [(1 - \delta) + \delta(1 - 2\varepsilon)^2 V^{MM} + \delta 2\varepsilon(1 - 2\varepsilon)(V^{DM} + V^{MD})] + \\ &\quad (\pi q + 1 - \pi) [-(1 - \delta) + \delta\varepsilon V^{MD}] + \pi(1 - p - q) [-2(1 - \delta) + \delta\varepsilon V^{MD}]; \\ V_D &= \pi p [\frac{15}{4}(1 - \delta) + \delta\varepsilon V^{DM}] + (\pi q + 1 - \pi)(0) + \pi(1 - p - q) [-(1 - \delta)]; \\ V_E &= \pi p [3(1 - \delta) + \delta\varepsilon V^{DM}] + (\pi q + 1 - \pi) [-(1 - \delta)] + \pi(1 - p - q) [-2(1 - \delta)]. \end{aligned}$$

of two states, the cooperative state (where players always randomize among C , D and E), and the defective state (where players play the stage game Nash equilibrium (D, D)). This is especially important for the belief-based approach since the main difficulty of extending belief-based strategies to general games comes from the fact that it becomes increasingly cumbersome to keep track of the evolution of players' beliefs when the stage game and the private monitoring structure become more complicated.

This construction, however, also imposes non-trivial restrictions on the specifications of the repeated game as for players to be willing to always randomize after good private histories, there should be no profitable one-shot deviation for each player. This incentive constraint (at each period after the good private histories, the continuation values of *playing each action in the support of the randomization and then conforming to* (ρ, π) should be identical) will be imposed jointly on stage game payoffs and the private monitoring structure. In the above, $(R - 1)$ is typically imposed on the discount factor and, as we have shown before, as long as a public randomization device is available, $(R - 1)$ usually poses no problem in applying modified belief-based strategies to general games. Restriction $(R - 2)$, on the other hand, usually imposes a non-trivial constraint on what type of games we can apply the modified belief-based construction. For example, for the original form of the game defined in **Figure 4** (Figure 2 in Sekiguchi (1997)), it is not clear how to apply the above method to construct a sequential equilibrium that is not the repetition of the stage Nash equilibrium. The reason is that if we use the above construction to support any payoff bigger than zero in a sequential equilibrium, the value of playing " D " and then conforming to (ρ, π) after good private histories should be positive. However, as long as the private monitoring structure has full support, the expected value of playing " D " is always negative in this game. So restriction $(R - 2)$ is violated. Thus for such games, we will have to design more complicated (possibly *time variant*) belief-based strategies to sustain cooperation.

4 Concluding Remarks

This paper introduces some form of ongoing randomizations, "*keep them guessing*" strategies, to extend the belief-based approach in Sekiguchi (1997) and Bhaskar and Obara (2002). We show that, using our modified belief-based strategies, efficiency can be approximated in repeated prisoners' dilemmas with asymmetric private monitoring, provided that observation errors are sufficiently small and a public randomization device is available for the players. The paper also provides an explicit example to show that the "*keep them guessing*" strategies can be potentially generalizable, under certain restrictions on the stage game payoffs and the private monitoring structure of the repeated game.

As mentioned in the introduction, our approach in this paper is a belief-based one where each player's beliefs of the other player's private histories are central to the analysis (hence the difficult statistical inferences in repeated games with private monitoring are explicitly addressed). This is in contrast to the belief-free approach where after each private history, a player's optimal continuation strategy is independent of the other player's private histories²⁵. This irrelevance of beliefs provides a drastic simplification of the analysis of repeated games with private monitoring. As a result, more general results have been obtained in the literature using this approach: *first*,

²⁵Note that our modified belief-based strategies are *different* from belief-free strategies: player l is indifferent between σ_l^M and σ_l^D *only* after a particular class of private histories of the form " Cc, \dots, Cc ". Player l 's beliefs are only fixed after such private histories and are still relevant for the optimality. If she sees a bad signal or her randomization realizes " D ", player l would then strictly prefer to player σ_l^D

the folk theorem for the repeated prisoners' dilemma with symmetric and asymmetric almost-perfect private monitoring has been obtained using the belief-free approach (Piccione (2002), Ely and Välimäki (2002)); *second*, Ely, Hörner and Olszewski (2005) characterize the entire set of equilibrium payoffs using "belief-free" sequential equilibria for two-player repeated games with private monitoring. In many games, the set of belief-free equilibrium payoffs is strictly larger than the convex hull of the stage game Nash equilibrium payoffs, but using only belief-free equilibria to obtain a folk theorem is almost always impossible for two-player repeated games even under almost-perfect private monitoring settings, as certain restrictions have to be imposed on the action profiles in constructing the belief-free strategies²⁶. The most general results (so far) on repeated games with almost-perfect private monitoring are probably provided by Hörner and Olszewski (2006) where the authors consider *block strategies*, which treat T^{th} -repeated stage game as a single stage game. Hörner and Olszewski (2006) show that, by considering block equilibria, it is possible to circumvent the restrictions of the belief-free approach on the action profiles and a folk theorem can be obtained for general N -player repeated games with almost perfect private monitoring. Although the strategies considered in Hörner and Olszewski (2006) are *not* belief-free, they contain the essential feature of the belief-free property: at the initial period of each T -block, each player is indifferent between the "reward" strategy and the "punishment" strategy and thus beliefs are irrelevant at the beginning of each block²⁷.

There are two potential criticisms of belief-free equilibria in the literature: first and most importantly, as belief-free strategies typically require a lot of randomizations for each player, *the coordination interpretation* common in equilibria of repeated games is not clear in the belief-free approach. The modified belief-based strategies constructed in this paper, on the other hand, have a clear coordination interpretation: the modified belief-based strategies are again *generalizations of trigger strategies* to the setting of private monitoring. These strategies are similar to trigger strategies in repeated games with perfect or imperfect public monitoring, which are constructed to provide enough incentives for the players to follow the equilibrium actions. The role of the randomizations along certain private histories is exactly the same as the role of first-period randomization (in symmetric private monitoring): to introduce *small* amounts of uncertainty so that players have proper incentives to punish bad signals, which is essential to deter defections.

The second criticism of belief-free equilibria is that it is not clear whether the belief-free equilibria are purifiable (Harsanyi (1973)). Bhaskar, Mailath and Morris (2006) investigate the purifiability of the mixed strategies used in constructing the belief-free equilibria in Ely and Välimäki (2002) in the perfect monitoring setting. They perturb the stage game (prisoners' dilemma) payoffs to allow *i.i.d.* private payoff shocks and find that almost always, all of the mixed strategy profiles are not the limit of one-period memory equilibrium strategy profiles of the perturbed game²⁸. As we have mentioned in *the fourth remark* after **Proposition 1**, the modified belief-based strategies we constructed seem to involve some unreasonable randomizations at only one information set (at $t = 0$ and after a private history of the form (Cc, \dots, Cc) , player l 's belief of h 's continuation strategy is identical, but player l is required to play differently). However, the modified belief-based strategies have *infinite history dependence* as each player's

²⁶A recent paper (Yamamoto (2006)) extends Ely, Hörner and Olszewski (2005) and fully characterizes the belief-free equilibrium payoffs in N -player repeated games with private monitoring.

²⁷Within each block, each player plays a complicated strategy where, almost always, positive probabilities are assigned to all actions of that player.

²⁸However, their non-purifiability arguments do not extend to the private monitoring case where a different method is necessary to show the non-purifiability results in the private monitoring settings.

beliefs of the opponent being cooperative are the essential part of the strategies and players beliefs are derived recursively from the entire previous private histories. For mixed strategies of infinitely repeated games that have infinite memories, no evidence on whether such strategies are purifiable or not has been provided in the literature. Indeed, future work should be directed towards investigating the purifiability of such mixed strategies.

Further extensions of belief-based strategies to more general repeated games are a difficult task. The main difficulty comes from the fact that it becomes increasingly cumbersome to keep track of the evolution of players' beliefs when the stage game and the private monitoring structure become more complicated. This paper presents one construction where keeping track of players' beliefs can be somewhat simplified by only considering two states for each player. And this construction can be potentially generalized to some class of repeated games with private monitoring. As belief-based strategies have a clear behavior interpretation, they are an important class of strategies to study for repeated games with private monitoring. We think that in the future, more work should be done in this line of research.

5 Appendix

5.1 Omitted Details in Section 2.1 – Section 2.2

The "coordination" game at $t = 0$ for the conventional belief-based strategies is again displayed as follows (note that, for example, V_i^{CD} denotes player i 's continuation value when i plays σ_i^C and the opponent plays σ_j^D):

	σ_h^C	σ_h^D
σ_l^C	V_l^{CC}, V_h^{CC}	V_l^{CD}, V_h^{DC}
σ_l^D	V_l^{DC}, V_h^{CD}	$0, 0$

FIGURE 2. The "coordination game" at $t = 0$.

As only partial continuation strategies are relevant to calculate equilibrium payoffs, we obtain:

$$\left\{ \begin{array}{l} V_l^{DD} = V_h^{DD} = 0; \\ V_l^{DC} = \frac{(1+G)(1-\delta)}{1-p_h\delta}, \quad V_h^{DC} = \frac{(1+G)(1-\delta)}{1-p_l\delta}; \\ V_l^{CD} = \frac{-L(1-\delta)}{1-p_l\delta}, \quad V_h^{CD} = \frac{-L(1-\delta)}{1-p_h\delta}; \\ V_l^{CC} = \frac{(1-\delta)+p_l\delta V_l^{DC}+p_h\delta V_l^{CD}}{1-(1-p_h-p_l-q)\delta}, \quad V_h^{CC} = \frac{(1-\delta)+p_h\delta V_h^{DC}+p_l\delta V_h^{CD}}{1-(1-p_h-p_l-q)\delta}. \end{array} \right.$$

And it is easy to verify that $V_i^{CC}, V_i^{DC} > V_i^{DD} = 0 > V_i^{CD}$ if $\delta > \frac{G}{1+G}$ and (p_h, p_l, q) is close to the origin.

The initial randomization probabilities can now be calculated as:

$$\left\{ \begin{array}{l} \hat{\pi}_l = \frac{-V_h^{CD}}{V_h^{CC}-V_h^{CD}-V_h^{DC}} = \frac{(1-\delta+p_h\delta+p_l\delta+q\delta)(1-p_l\delta)L}{(1-p_l\delta)(1-p_h\delta)+(1-\delta+\delta p_h+\delta q)(1-p_l\delta)L-(1-\delta+\delta p_l+\delta q)(1-p_h\delta)(1+G)}; \\ \hat{\pi}_h = \frac{-V_l^{CD}}{V_l^{CC}-V_l^{CD}-V_l^{DC}} = \frac{(1-\delta+p_h\delta+p_l\delta+q\delta)(1-p_h\delta)L}{(1-p_l\delta)(1-p_h\delta)+(1-\delta+\delta p_l+\delta q)(1-p_h\delta)L-(1-\delta+\delta p_h+\delta q)(1-p_l\delta)(1+G)}. \end{array} \right.$$

And we have $\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}, \delta \rightarrow \frac{G}{1+G}} \hat{\pi}_i = 1$, and $\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}, \delta \rightarrow 1} \hat{\pi}_i = 0, \forall i \in \{l, h\}$.

Now, to show the result of **Lemma 1** ((5) and (6)) in more detail, note first that the belief operator $\chi_h^{Cc}(\mu)$ has a fixed point $\chi_h^* = \frac{1-2p_h-p_l-2q}{1-2p_h-2q} \in (0, 1)$, when p_h, p_l and q are small. To see what happens to player h 's belief of l being in the *cooperative* state after a private history

of the form " Cc, \dots, Cc, Cd ", we can calculate $\chi_h^{Cd}(\chi_h^*)$, which can be made arbitrarily close to player h 's belief after " Cc, \dots, Cc, Cd " with an arbitrarily long sequence of " Cc "s:

$$\chi_h(Cc, \dots, Cc, Cd) \approx \chi_h^{Cd}(\mu) \approx \chi_h^{Cd}(\chi_h^*) = \frac{p_h}{(p_h + q) + \frac{(1-p_h-q)p_l}{1-2p_h-p_l-2q}}.$$

Now, fixing $h > 0$ and letting l (and thus q) go to zero, we have $\lim_{l \rightarrow 0} \chi_h^{Cd}(\chi_h^*) = 1$, which means that if private signals are such that p_h is significantly larger than p_l , player h 's belief of l being in the "cooperative" state can be very large after a private history " $\underbrace{Cc, \dots, Cc, Cd}_n$ " when n is

very large. This completes the description of result (5). On the other hand, player l 's initial randomization probability (note that $\hat{\pi}_l$, defined in (1), is also the threshold where player h 's behavior changes, by the criterion function defined in (4)) when p_l and q become arbitrarily small can be calculated as (result (6)):

$$\lim_{p_l, q \rightarrow 0} \hat{\pi}_l = \frac{(1 - \delta + \delta p_h)(1 - \delta p_h)L}{1 + (1 - \delta)L - \frac{(1+G)}{(1-\delta p_h)}(1 - \delta + \delta p_h)} < 1,$$

if $\delta > \frac{G}{1+G}$ and p_h is small. Hence, as long as the " n " in history " $\underbrace{Cc, \dots, Cc, Cd}_n$ " is large enough

and p_l, q are sufficiently smaller than p_h , we will have $\chi_h^{Cd}(\mu) > \hat{\pi}_l$ (Note that as the belief operator $\chi_h^{Cd}(\mu)$ is strictly increasing in μ . Hence, $\chi_h^{Cd}(\chi_h^*(Cc))$ is defined as the *supremum* of the belief operator $\chi_h^{Cd}(\mu)$, since after a long sequence of " Cc "s, h 's belief can be arbitrarily close to $\chi_h^*(Cc)$, and the specific condition under which the result in **Lemma 1** appears is " $\chi_h^{Cd}(\chi_h^*) \geq \hat{\pi}_l$ ".) Thus the conventional belief-based strategy profiles in Sekiguchi (1997) and Bhaskar and Obara (2002) are not Nash equilibria when $p_h \gg p_l, q$ as player h now has no incentive to switch to defection after a bad signal on the equilibrium path.

5.2 Omitted Details in Section 2.3

First, the initial randomization probabilities for the modified belief-based strategies are derived as follows:

$$\begin{array}{c} \sigma_l^M \\ \sigma_l^D \end{array} \begin{array}{|c|c|} \hline \begin{array}{c} \sigma_h^C \\ V_l^{MC}, V_h^{CM} \end{array} & \begin{array}{c} \sigma_h^D \\ V_l^{MD}, V_h^{DM} \end{array} \\ \hline \begin{array}{c} V_l^{DC}, V_h^{CD} \end{array} & 0, 0 \\ \hline \end{array}$$

The "coordination game" at $t = 0$.

As only the partial continuation strategies are relevant in evaluating the expected values of the equilibrium, we can thus calculate these values as:

- Player h :

$$\begin{cases} V_h^{CM} = \frac{(1-\delta)(p+pL-L) + \delta p_h p V_h^{DM} + \delta [pp_l + (1-p)(p_h+q)] V_h^{CD}}{1 - \delta p(1-p_h-p_l-q)}; \\ V_h^{DM} = (1-\delta)(1+G)p + \delta(p_l+q)p V_h^{DM} = \frac{(1-\delta)p(1+G)}{1-p\delta(p_l+q)}; \\ V_h^{CD} = -L(1-\delta) + \delta(p_h+q)V_h^{CD} = \frac{-(1-\delta)L}{1-\delta(p_h+q)}; \\ V_h^{DD} = 0. \end{cases}$$

- Player l :

$$\begin{cases} V_l^{MC} = \frac{(1-\delta)(1+G-pG) + \delta p_h p V_l^{MD} + \delta [p p_l + (1-p)(p_h+q)] V_l^{DC}}{1-\delta p(1-p_h-p_l-q)}; \\ V_l^{MD} = -pL(1-\delta) + \delta(p_l+q)p V_l^{MD} = \frac{-(1-\delta)pL}{1-p\delta(p_l+q)}; \\ V_l^{DC} = (1-\delta)(1+G) + \delta(p_h+q)V_l^{DC} = \frac{(1-\delta)(1+G)}{1-\delta(p_h+q)}; \\ V_l^{DD} = 0. \end{cases}$$

The probabilities of the initial randomizations can thus be determined as:

$$\begin{cases} \pi_l = \frac{-V_h^{CD}}{V_h^{CM} - V_h^{CD} - V_h^{DM}}; \\ \pi_h = \frac{-V_l^{MD}}{V_l^{MC} - V_l^{MD} - V_l^{DC}}. \end{cases}$$

Proof of Lemma 2

Firstly, simplifying (9) gives the following condition:

$$\begin{aligned} F(\delta; p_h, p_l, q) &\equiv (1-p_h-2p_l-2q) \left[\begin{array}{c} (1-2p_h-2q)\delta \\ -(1-\delta+\delta p_h+\delta q)G \end{array} \right] - p_h L(1-\delta p_h-\delta q) \\ &= 0. \end{aligned}$$

Note that this restriction is imposed on δ and that there is a $\delta^* \in (0, 1)$ satisfying $F(\delta; p_h, p_l, q) = 0$ when p_h, p_l and q are small enough (**Lemma 2**). In addition, δ^* *only* depends on (p_h, p_l, q) , but does *not* depend on the probability p (to be defined). The underlying reason is that player h *only* randomizes in the initial period, and given player h 's strategy and $\pi_h = \tilde{\chi}_l^{Cc}$, player l is always indifferent between choosing C and D in each period after a good private history " Cc, \dots, Cc ". This gives us some freedom in choosing p (see the next restriction) and it will also simplify the analysis on p . However, when both players adopt partial continuation strategies $\sigma = \{\sigma^M, \sigma^D\}$, δ^* would also depend on p from restriction $\pi_h = \tilde{\chi}_l^{Cc}$. We will see this more clearly in **Section 3**, where we extend the modified belief-based strategies to some other game (there, *both* players have to randomize more than once on the equilibrium path).

Secondly, simplifying (11) yields:

$$\begin{aligned} H(p; \delta, p_h, p_l, q) &\equiv (p + \delta p + 2\delta p^2 p_h + 2\delta p^2 q - 2\delta p p_h - 2\delta p q - \delta p p_l - \delta p^2 - 1) X \\ &\quad - (p - \delta p^2 + \delta p^2 p_l + \delta p^2 q) Y + (p^2 + p^2 L - pL) \\ &= 0, \end{aligned}$$

where $X = \frac{L}{1-\delta(p_h+q)}$, $Y = \frac{p(1+G)}{1-p\delta(p_l+q)}$.

Proof.

- δ^* :

From equation $F(\delta; p_h, p_l, q) = 0$, we have that

$$\frac{\partial F}{\partial \delta} = (1-p_h-2p_l-2q)[1-2p_h-2q+(1-p_h-q)G] + p_h^2 L + p_h q L.$$

Since $\frac{\partial F}{\partial \delta}|_{(p_h, p_l, q)=\mathbf{0}} = (1+G) > 0$, by the implicit function theorem, there is an open set U' of (p_h, p_l, q) around the origin $(0, 0, 0)$ such that $F(\delta; p_h, p_l, q) = 0$ implicitly defines δ as a unique C^1 function of (p_h, p_l, q) , $\delta^*(p_h, p_l, q)$. From $F(\delta; p_h, p_l, q) = 0$, we also

have $\delta^*(0, 0, 0) = \frac{G}{1+G}$. Apply the result from the implicit function theorem again and we have that in the open set U' of (p_h, p_l, q) , the partial differentiations of $\delta^*(p_h, p_l, q)$ can be calculated as follows:

$$\frac{\partial \delta}{\partial p_h} = -\frac{\frac{\partial F}{\partial p_h}}{\frac{\partial F}{\partial \delta}}; \quad \frac{\partial \delta}{\partial p_l} = -\frac{\frac{\partial F}{\partial p_l}}{\frac{\partial F}{\partial \delta}}; \quad \frac{\partial \delta}{\partial q} = -\frac{\frac{\partial F}{\partial q}}{\frac{\partial F}{\partial \delta}}.$$

Simple calculations yield:

$$\begin{cases} \frac{\partial F}{\partial p_h}|_{(p_h, p_l, q)=\mathbf{0}} = -L - \frac{G^2+2G}{1+G} < 0; \\ \frac{\partial F}{\partial p_l}|_{(p_h, p_l, q)=\mathbf{0}} = 0; \\ \frac{\partial F}{\partial q}|_{(p_h, p_l, q)=\mathbf{0}} = -\frac{2G+G^2}{1+G} < 0. \end{cases}$$

The above result implies that there is an open set U'' near the origin where $p_h, p_l, q > 0$ in U'' and (1) $\frac{\partial F}{\partial p_h}, \frac{\partial F}{\partial p_l}, \frac{\partial F}{\partial q} < 0$ ²⁹ in U'' (with $p_h > p_l$); (2) $\delta^*(p_h, p_l, q) \in \left(\frac{G}{1+G}, 1\right)$ in U'' (with $p_h > p_l$). Taking \mathbf{U} as $\mathbf{U} = U' \cap U''$, we have then proved part 1.

- p :

Now suppose that $(p_h, p_l, q) \in \mathbf{U}$, and equation $H(p; \delta^*, p_h, p_l, q) = 0$, where δ^* only depends on (p_h, p_l, q) in \mathbf{U} , can be written as $\widehat{H}(p; p_h, p_l, q) = 0$. This implies, using $\delta^*|_{(p_h, p_l, q)=\mathbf{0}} = \frac{G}{1+G}$,

$$\widehat{H}(p; 0, 0, 0) = \left(1 + G + \frac{L}{1+G}\right)p^2 - \left(1 + G - \frac{GL}{1+G}\right)p - L = 0,$$

from which we can solve that $p|_{(p_h, p_l, q)=\mathbf{0}} = 1$ (or $p|_{(p_h, p_l, q)=\mathbf{0}} < 0$). Discarding the negative solution, we have $\widehat{H}(1; 0, 0, 0) = 0$, which in turn implies that

$$\frac{\partial \widehat{H}}{\partial p}|_{(p_h, p_l, q)=\mathbf{0}} = (1+G) + \left(1 + \frac{1}{1+G}\right)L > 0.$$

Hence, again, by the implicit function theorem, there exists some open set V' near $(0, 0, 0)$ such that $H(p; \delta, p_h, p_l, q) = 0$ implicitly defines p as a unique C^1 function of (p_h, p_l, q) . To determine the partial differentiations, we can easily verify that $\frac{\partial \widehat{H}}{\partial p_h}|_{(p_h, p_l, q)=\mathbf{0}} > 0$ and $\frac{\partial \widehat{H}}{\partial q}|_{(p_h, p_l, q)=\mathbf{0}} > 0$, which by the implicit function theorem we have $\frac{\partial p}{\partial p_h}$ and $\frac{\partial p}{\partial q}$ are negative in the open set V' . Sets \mathbf{U} and V' determine that there another open set \mathbf{V} such that result 2 holds for $(p_h, p_l, q) \in \mathbf{V}$. Note that $\frac{\partial \widehat{H}}{\partial p_l}|_{(p_h, p_l, q)=\mathbf{0}} < 0$. However, by comparing $\frac{\partial \widehat{H}}{\partial p_l}|_{(p_h, p_l, q)=\mathbf{0}}$ and $\frac{\partial \widehat{H}}{\partial p_h}|_{(p_h, p_l, q)=\mathbf{0}}$, we have that

$$\left| \frac{\partial \widehat{H}}{\partial p_h}|_{(p_h, p_l, q)=\mathbf{0}} \right| = L + \frac{G^2 + 2G}{1+G} > \left| \frac{\partial \widehat{H}}{\partial p_l}|_{(p_h, p_l, q)=\mathbf{0}} \right| = \frac{GL}{1+G} + \frac{2G + G^2}{1+G},$$

and this condition ensures that in the open set \mathbf{V} , the probability p can be restricted to be in $(0, 1)$.

■

²⁹The result " $\frac{\partial F}{\partial l}|_{(p_h, p_l, q)=\mathbf{0}} = 0$ " does not pose any problem for us. Note U'' does *not* include the origin and we are concerned with the case where $p_h > p_l$. It's easy to see that by (1) $\frac{\partial \delta}{\partial p_h} > 0$ and $\frac{\partial \delta}{\partial q} > 0$ in U' , and (2) From $\frac{\partial F}{\partial p_l} = -2[\delta(1-2p_h) - (1-\delta+\delta p_h)G]$, we have that $\frac{\partial^2 F}{\partial p_l \partial p_h} < 0$ at the origin. This implies that $\frac{\partial F}{\partial p_l} < 0$ whenever p_h is small and positive.

Proof of Lemma 3

Proof. Part 1 and part 3 are straightforward to verify.

To show part 2, note that $\pi_l < \chi_h^*$ is equivalent to " $(1 - p_h - q)(\pi_l - \pi_l^2) > (p_h + q + p_l\pi_l) - (p_h + q)\pi_l^2$ ". Now it is easy to see that as $(1 - p_h - q)\pi_l > (p_h + q)(1 + \pi_l)$ when p_h and q are small enough, the following inequality holds strictly:

$$(1 - p_h - q)(\pi_l - \pi_l^2) > (p_h + q) - (p_h + q)\pi_l^2.$$

Accordingly, the original inequality " $(1 - p_h - q)(\pi_l - \pi_l^2) > (p_h + q + p_l\pi_l) - (p_h + q)\pi_l^2$ " is true if we choose p_l to be sufficiently small. Thus there is an open set $\mathbf{W} \subset \mathbf{U} \cap \mathbf{V}$ such that $\pi_l < \chi_h^*$ for all $(p_h, p_l, q) \in \mathbf{W}$. ■

Proof of Lemma 4

Proof.

- For player h , we need to show that:

1. " $\pi_l \leq \mu < \chi_h^*$ implies **(A)** $\chi_h^{Cc}(\mu) > \pi_l$ and **(B)** $\chi_h^{Cd}(\mu) < \pi_l$ ".

For **(1)**, to verify " $\chi_h^{Cc}(\mu) > \pi_l$ ", it suffices to show " $\chi_h^{Cc}(\pi_l) > \pi_l$ " as by part 3 in **Lemma 3**, $\frac{d\chi_h^{Cc}(\mu)}{d\mu} > 0$. But " $\chi_h^{Cc}(\pi_l) > \pi_l$ " is equivalent to the following,

$$\begin{aligned} (1 - p_h - p_l - q)\pi_l &> (1 - 2p_h - 2q)\pi_l^2 + (p_h + q) \\ &\Leftrightarrow \\ (1 - p_h - q)(\pi_l - \pi_l^2) &> p_l\pi_l + (p_h + q) - (p_h + q)\pi_l^2. \end{aligned}$$

Now by the proof of part 2 in **Lemma 3**, $\pi_l < \chi_h^*$ is equivalent to " $(1 - p_h - q)(\pi_l - \pi_l^2) > (p_h + q + p_l) - (p_h + q)\pi_l^2$ ". Hence, the last inequality above is reduced to " $p_l > p_l\pi_l$ ", which is true since " $(p_h, p_l, q) \in \mathbf{W} \subset \mathbf{U} \cap \mathbf{V}$ " implies that " $\pi_l < 1$ ".

For **(2)**, note that $\frac{d\chi_h^{Cd}(\mu)}{d\mu} > 0$ by **Lemma 3**. Hence it suffices to show " $\chi_h^{Cd}(\chi_h^*) < \pi_l$ ", which is equivalent to " $(1 - 2p_h - 2q)\chi_h^*\pi_l + p_h\chi_h^* < 1 - p_h - q$ ". But we have

$$\begin{aligned} (1 - 2p_h - 2q)\chi_h^*\pi_l + p_h\chi_h^* &< (1 - 2p_h - 2q)\chi_h^* + p_h\chi_h^* \\ &< (1 - p_h - q), \end{aligned}$$

where the first inequality comes from the fact that " $(p_h, p_l, q) \in \mathbf{W}$ " implies " $\pi_l < 1$ " and the second inequality comes from " $\chi_h^* < 1$ " (**Lemma 3**).

2. " $0 < \mu < \pi_l$ implies **(C)** $\chi_h^{Dc}(\mu) < \pi_l$ and **(D)** $\chi_h^{Dd}(\mu) < \pi_l$ ".

(C) is equivalent to " $(1 - 2p_h - 2q)\mu\pi_l + p_h + q > p_l\mu$ " for $0 < \mu < \pi_l$. This is obviously true since " $p_h + q > p_l > p_l\mu$ " and " $p_h + q < \frac{1}{2}$ ". And **(D)** can be proved similarly.

- For player l , her partial continuation strategy specifies that she play " $\pi_l \circ \sigma_l^M + (1 - \pi_l) \circ \sigma_l^D$ " at $t = 0$ and she play σ_l^M whenever she sees " Cc, \dots, Cc ":

1. At $t = 0$, given that player h is playing " $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$ ", player l 's belief is $\mu_l = \pi_h = \frac{1-p_h-2p_l-2q}{1-2p_l-2q}$, and player l plays " $\pi_l \circ \sigma_l^M + (1 - \pi_l) \circ \sigma_l^D$ ". So at $t = 0$, player l plays " $\pi_l^2 \circ C + (1 - \pi_l^2) \circ D$ ", as we imposed the condition $p = \pi_l$. For $t \geq 1$, given that player l 's private history is " Cc, \dots, Cc ", her belief that player h being in the cooperative state is *always* π_h , and player l 's continuation strategy will thus be " $\pi_l \circ \sigma_l^C + (1 - \pi_l) \circ \sigma_l^D$ ". Player l 's partial continuation strategy is assumed to be σ_l^M , or at each period with a private history " Cc, \dots, Cc ", player l plays " $\pi_l \circ C + (1 - \pi_l) \circ D$ ".
2. When player l sees " Cd " after a history " Cc, \dots, Cc ", she is required to play σ_l^D . Or we need to show " $\mu = \pi_h = \chi_l^*$ implies $\chi_l^{Cd}(\mu) < \pi_h$ ". But this is equivalent to " $p_l + (1 - 2p_l - 2p) \pi_h < 1 - p_l - p$ ", which is true since $\pi_h = \frac{1-p_h-2p_l-2q}{1-2p_l-2q} < 1$.
3. Finally, we need to show that " $0 < \mu \leq \pi_h$ implies **(E)** $\chi_l^{Dc}(\mu) < \pi_h$ and **(F)** $\chi_l^{Dd}(\mu) < \pi_h$ ". Since by part 3 in **Lemma 3**, $\frac{d\chi_l^{Dc}(\mu)}{d\mu}, \frac{d\chi_l^{Dd}(\mu)}{d\mu} > 0$, it suffices to show that $\chi_l^{Dc}(\pi_h) < \pi_h$ and $\chi_l^{Dd}(\pi_h) < \pi_h$. These two inequalities reduce to

$$\begin{cases} p_h < \pi_h (1 - 2p_l - 2q) + p_l + q; \\ q + (1 - 2p_l - 2q) \pi_h < (1 - p_l - q). \end{cases}$$

It is easy to see that there exists an open set $\widehat{\mathbf{W}} \subset \mathbb{R}^3$ near the origin such that above inequalities are true for all $(p_h, p_l, q) \in \widehat{\mathbf{W}}$. For the rest of the paper, denote $\mathbf{Z} = \widehat{\mathbf{W}} \cap \mathbf{W}$.

By the arguments shown above, the modified belief-based strategies defined in (15) and (16) are realization equivalent to σ_l and σ_h , respectively. ■

Proof of Proposition 1

To accommodate the strategy profile in this paper, we first define players' criterion functions as follows:

$$\begin{cases} \Delta V_h(\mu; \delta^*, (p_h, p_l, q)) = \mu (V_h^{CM} - V_h^{DM}) - (1 - \mu) (V_h^{DD} - V_h^{DM}); \\ \Delta V_l(\mu; \delta^*, (p_h, p_l, q)) = \mu (V_l^{MC} - V_l^{DC}) - (1 - \mu) (V_l^{DD} - V_l^{DC}). \end{cases}$$

Note again that there is a *unique* vector $\boldsymbol{\mu}$, such that $\Delta V_h(\mu; \delta^*, (p_h, p_l, q)) = \Delta V_l(\mu; \delta^*, (p_h, p_l, q)) = 0$ by the linearity of these two functions in μ .

Proof. The objective is to show that there is no profitable one-shot deviation for any player at any private information set.

- Player h :

Case 1. $\mu = \pi_l$, by the definition of π_l , player h is indifferent between σ_h^C and σ_h^D . By construction, one-shot deviations are not profitable.

Case 2. If in the middle of the game, $\mu > \pi_l$, player h is required to play σ_h^C . A one-shot deviation is to play D in this period and then continue with (ρ_h, π_l) . The deviation gain depends on the belief operators $\chi_h^{Dc}(\mu)$ and $\chi_h^{Dd}(\mu)$. First, note that for all

$\mu \in [0, \chi_h^*]$, we have $\chi_h^{Dc}(\mu), \chi_h^{Dd}(\mu) < \pi_l$. The reason is the following: by part 3 in **Lemma 3**, it suffices to show $\chi_h^{Dc}(\chi_h^*), \chi_h^{Dd}(\chi_h^*) < \pi_l$, which are equivalent to

$$\begin{cases} \chi_h^* p_l < \chi_h^* \pi_l (1 - 2p_h - 2q) + p_h + q; \\ q\chi_h^* + (1 - 2p_h - 2q)\chi_h^* \pi_l < 1 - p_h - q. \end{cases}$$

These are true since $p_h + q > p_l > \chi_h^* l$ and $(p_h, p_l, q) \in \mathbf{Z}$ (thus $\chi_h^*, \pi_l < 1$). Hence, we only³⁰ need to consider the situation of $\chi_h^{Dc}(\mu), \chi_h^{Dd}(\mu) < \pi_l$.

Now, if $\chi_h^{Dc}(\mu), \chi_h^{Dd}(\mu) < \pi_l$, the one-shot deviation is equivalent to σ_h^D . As the criterion function $\Delta V_h(\mu; \delta^*, (p_h, p_l, q)) > 0$ for $\mu > \pi_l$, or σ_h^C is better than σ_h^D , this one-shot deviation is not profitable.

Case 3. Now suppose $\mu < \pi_l$. The one-shot deviation for h is to play C in this period and continue with (ρ_h, π_l) . Once h has played C , (ρ_h, π_l) specifies that h play σ_h^C if $\chi_h^{Cc}(\mu) \geq \pi_l$ and that h play σ_h^D if $\chi_h^{Cc}(\mu) < \pi_l$ (player h should also play σ_h^D if signal "d" is observed for h). If the private state $\chi_h^{Cc}(\mu) \geq \pi_l$ is reached, the payoff difference between the equilibrium continuation strategy (σ_h^D) and the one-shot deviation is the payoff difference between σ_h^C and σ_h^D when $\mu < \pi_l$. Again, as " $\mu < \pi_l$ implying $\Delta V_h(\mu; \delta^*, (p_h, p_l, q)) < 0$ ", σ_h^D is hence strictly better and this one-shot deviation is strictly worse for h . On the other hand, if we are in the other situation ($\chi_h^{Cc}(\mu) < \pi_l$ or "d" is observed), the payoff difference between the one-shot deviation (which resembles σ_h^C , with the exception that h plays σ_h^D after "C" and a good signal "c") and σ_h^D is the following³¹:

$$\begin{aligned} \Delta \tilde{V}_h(\mu) &= \Delta V_h(\mu; \delta^*, (p_h, p_l, q)) \\ &\quad - \delta^* [\mu \pi_l (1 - p_h - q) + (1 - \mu \pi_l) (p_h + q)] \Delta V_h(\chi_h^{Cc}(\mu); \delta^*, (p_h, p_l, q)). \end{aligned}$$

First, it is easy to show that $\chi_h^{Cc}(\mu) > \mu$ (**Lemma 4**). Now, if $\pi_l > \chi_h^{Cc}(\mu) > \mu$, then we have $0 > \Delta V_h(\chi_h^{Cc}(\mu); \delta^*, (p_h, p_l, q)) > \Delta V_h(\mu; \delta^*, (p_h, p_l, q))$, which implies that $\Delta \tilde{V}_h(\mu) < 0$, since $\delta^* [\mu \pi_l (1 - p_h - q) + (1 - \mu \pi_l) (p_h + q)] \in (0, 1)$. Therefore, the one-shot deviation of player h is not profitable.

- Player l :

First, note that by the construction of the strategy, player l 's belief is restricted to the interval $[0, \pi_h]$. Therefore, we only consider two cases: $\mu = \pi_h$ and $\mu < \pi_h$.

Case 1. $\mu = \pi_h$.

At $t = 0$, or whenever player l sees a history like "Cc, ..., Cc", player l 's belief is $\mu = \pi_h$, or l 's belief of h 's continuation strategy is always $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$.

³⁰Note the difference between the proof in **Case 2** and the corresponding part of proof in Bhaskar and Obara (2002): there is only one situation to consider in this paper (while there are two cases in B&O (2002)). The underlying reason comes from the definitions of the belief operators $\chi_h^{Dc}(\mu)$ and $\chi_h^{Dd}(\mu)$ in this paper.

³¹Note that

$$\begin{aligned} \Pr(s_h = c | A_h = C, \mu) &= \mu \pi_l \Pr(s_h = c | Cc) + (1 - \mu \pi_l) \Pr(s_h = c | Cd) \\ &= \mu \pi_l (1 - p_h - q) + (1 - \mu \pi_l) (p_h + q). \end{aligned}$$

The value of playing "C" and the value of playing "D" to player l are (given player h 's continuation strategy is σ_h^C):

$$\begin{aligned} V_l^C &= (1 - \delta) + \delta [(1 - p_h - p_l - q) V_l^{MC} + p_h V_l^{MD} + p_l V_l^{DC}], \\ V_l^D &= (1 - \delta)(1 + G) + \delta (p_h + q) V_l^{DC}, \end{aligned}$$

while the corresponding values to player l when player h 's continuation strategy is σ_h^D are:

$$\begin{aligned} \widehat{V}_l^C &= -L(1 - \delta) + \delta (p_l + q) V_l^{MD}, \\ \widehat{V}_l^D &= 0. \end{aligned}$$

Given that l believes "player h 's continuation strategy is $\pi_h \circ \sigma_h^C + (1 - \pi_h) \circ \sigma_h^D$ ", it is easy to show that (recall that condition (9) is the restriction that δ is fixed to a certain value):

$$(9) \implies \pi_h V_l^C + (1 - \pi_h) \widehat{V}_l^C = \pi_h V_l^D,$$

or no one-shot deviation from σ_l^M is profitable for player l if (9) is true.

Case 2. $\mu < \pi_h$. The one-shot deviation for l is to play C in this period and continue with (ρ_l, π_h) . Once l has played C, (ρ_l, π_h) specifies that l play σ_l^M if $\chi_h^{Cc}(\mu) = \pi_l$ and l play σ_l^D if $\chi_l^{Cc}(\mu) < \pi_h$, or if signal "d" is observed for l . Now, if it is the case that $\chi_h^{Cc}(\mu) = \pi_l$, then the one-shot deviation corresponds to " σ_l^M ", but by the criterion function we know that " σ_l^D " is always strictly preferred to " σ_l^M ".

Now consider the second case where $\chi_l^{Cc}(\mu) < \pi_h$. The one-shot deviation resembles the (partial) continuation strategy " σ_l^M ", except at the information set "Cc", where l plays " σ_l^D " instead of " σ_l^M ". Denote the difference in values between playing " σ_l^D " and playing this one-shot deviation to be $\Delta \widetilde{V}(\mu)$ and we have (let $\mu' = \chi_l^{Cc}(\mu)$):

$$\Delta \widetilde{V}(\mu) = \Delta V(\mu) - \delta [(1 - p_h - q) \mu \pi_l + (p_h + q)(1 - \mu \pi_l)] \Delta V(\mu').$$

Now, since $\Delta V(\pi_h) = 0$ and $\mu < \mu' = \chi_l^{Cc}(\mu) < \pi$, it is immediately evident that $\Delta \widetilde{V}(\mu) > 0$, or the one-shot deviation is not profitable.

By the above arguments for player l and player h and **Lemma 4**, the belief-based strategy profile $((\rho_l, \pi_h), (\rho_h, \pi_l))$ is a sequential equilibrium for the repeated games with asymmetric private monitoring. ■

Proof of Proposition 2

Proof. By **Proposition 1**, the modified belief-based strategy profile $((\rho_l, \pi_h), (\rho_h, \pi_l))$ is a sequential equilibrium provided that (p_h, p_l, q) is close enough to the origin. The result in **Proposition 1** continues to hold for the situation where players face an even more accurate monitoring structure (that is, for vectors of (p_h, p_l, q) that are even closer to the origin). Applying **Lemma 5**, together with the fact that $\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} V_h^{CM} = 1$ and $\lim_{(p_h, p_l, q) \rightarrow \mathbf{0}} V_l^{MC} = 1$, we know that for any $\delta \geq \delta^*$ (δ^* is defined in **Lemma 2**), there exists a sequential equilibrium where players obtain payoffs arbitrarily close to (1, 1), provided that p_h, p_l and q are arbitrarily close to 0. ■

5.3 Omitted Details in Section 3

First, the values of the equilibrium to the players in **Figure 7** can be calculated as

$$\begin{cases} V^{MM} = \frac{(1-\delta)(5p+2q-\frac{5}{4}pq-2p^2-2)+\delta\varepsilon p(1+p-2p\varepsilon)V^{MD}+\delta\varepsilon p(1+p-4p\varepsilon)V^{DM}}{1-\delta p^2(1-2\varepsilon)^2}; \\ V^{MD} = \frac{(1-\delta)(q-1)}{1-\delta p\varepsilon}; \quad V^{DM} = \frac{(1-\delta)(\frac{15}{4}p+q-1)}{1-\delta p\varepsilon}; \\ V^{DD} = 0. \end{cases}$$

We now present the belief dynamics of the players when they employ the modified belief-based strategies. Using Bayes' rule, we have³²:

$$\begin{aligned} \chi^{Cc}(\mu) &= \frac{(1-2\varepsilon)^2 p\mu}{(1-2\varepsilon)\mu p + \varepsilon(1-\mu p)}; \\ \chi^{Cd}(\mu) &= \frac{\varepsilon(1-2\varepsilon)p\mu}{(1-2\varepsilon) - (1-3\varepsilon)(1-q)\mu}; \\ \chi^{Ce}(\mu) &= \frac{\varepsilon(1-2\varepsilon)p\mu}{\varepsilon + (1-3\varepsilon)(1-p-q)\mu}; \\ \chi^{Dd}(\mu) &= \frac{\varepsilon(1-2\varepsilon)p\mu}{\varepsilon + (1-3\varepsilon)(1-\mu+q\mu)}; \\ \chi^{De}(\mu) &= \frac{\varepsilon^2 p\mu}{\varepsilon + (1-3\varepsilon)(1-p-q)\mu}; \\ \chi^{Ec}(\mu) &= \frac{\varepsilon(1-2\varepsilon)p\mu}{\varepsilon + (1-3\varepsilon)p\mu}; \\ \chi^{Ed}(\mu) &= \frac{\varepsilon^2 p\mu}{\varepsilon + (1-3\varepsilon)(1-\mu+q\mu)}; \\ \chi^{Ee}(\mu) &= \frac{\varepsilon^2 p\mu}{\varepsilon + (1-3\varepsilon)(1-p-q)\mu}. \end{aligned}$$

Now it is easy to verify that the full belief-based strategy defined in (30) is realization equivalent to σ_i defined in (29), as shown by the following facts:

- Restriction $(R-1)$, $\pi = \chi_{Cc}^*$, ensures that a player's belief of the other being cooperative is always χ_{Cc}^* after the good histories of the form " Cc, \dots, Cc ";
- By the fact that $\frac{d\chi^{Cd}(\mu)}{d\mu}, \frac{d\chi^{Ce}(\mu)}{d\mu} > 0$, it is sufficient to verify $\chi^{Cd}(\pi), \chi^{Ce}(\pi) < \pi$, which hold for ε positive and small (The argument uses " $\pi, p, q \in (0, 1)$ when ε is small enough", which will be verified shortly);

³²For example:

$$\begin{aligned} \chi^{Cd}(\mu) &= \frac{\Pr(dc|CC)p\mu}{p\mu[\Pr(dc|CC) + \Pr(dd, e|CC)] + (1-\mu+\mu q)[\Pr(dc|CD) + \Pr(dd, e|CD)] + (1-p-q)\mu[\Pr(dc|CE) + \Pr(dd, e|CE)]}; \\ \chi^{Dc}(\mu) &= \frac{\Pr(cc|DC)p\mu}{p\mu[\Pr(cc|DC) + \Pr(cd, e|DC)] + (1-\mu+\mu q)[\Pr(cc|DD) + \Pr(cd, e|DD)] + (1-p-q)\mu[\Pr(cc|DE) + \Pr(cd, e|DE)]}; \\ \chi^{Ee}(\mu) &= \frac{\Pr(ec|EC)p\mu}{p\mu[\Pr(ec|EC) + \Pr(ed, e|EC)] + (1-\mu+\mu q)[\Pr(ec|ED) + \Pr(ed, e|ED)] + (1-p-q)\mu[\Pr(ec|EE) + \Pr(ed, e|EE)]}. \end{aligned}$$

- It is immediate to verify that

$$\frac{d\chi^{Dc}(\mu)}{d\mu}, \frac{d\chi^{Dd}(\mu)}{d\mu}, \frac{d\chi^{De}(\mu)}{d\mu} > 0,$$

and

$$\frac{d\chi^{Ec}(\mu)}{d\mu}, \frac{d\chi^{Ed}(\mu)}{d\mu}, \frac{d\chi^{Ee}(\mu)}{d\mu} > 0.$$

Thus, it suffices to show that $\chi^{Dc}(\pi), \chi^{Dd}(\pi), \chi^{De}(\pi), \chi^{Ec}(\pi), \chi^{Ed}(\pi), \chi^{Ee}(\pi)$ are all less than π , which is obviously true as long as " $\pi, p, q \in (0, 1)$ ".

Now we verify the claim that "there exists $\bar{\varepsilon} > 0$, such that for all $\varepsilon \in (0, \bar{\varepsilon})$, p, q, δ and π are C^1 functions of ε and in the neighborhood $(0, \bar{\varepsilon})$, $\frac{d\delta}{d\varepsilon}, \frac{dp}{d\varepsilon}, \frac{d\pi}{d\varepsilon} > 0$ and $\frac{dq}{d\varepsilon} < 0$. In addition, we have

$$\lim_{\varepsilon \rightarrow 0} \delta = \frac{10}{11}, \lim_{\varepsilon \rightarrow 0} p = \frac{4}{5}, \lim_{\varepsilon \rightarrow 0} q = \frac{1}{5}, \lim_{\varepsilon \rightarrow 0} \pi = 1."$$

The basic argument is again the implicit function theorem.

First, by restriction $(R - 1)$ and " $V_D = V_E$ ", we have

$$\frac{4}{5}\pi p = 1 \text{ and } p = \frac{4 - 7\varepsilon}{5(1 - 2\varepsilon)^2},$$

from which we know p is a differentiable function of ε and $p \in (0, 1)$, $\frac{dp}{d\varepsilon} > 0$ if ε is small. Also, $\lim_{\varepsilon \rightarrow 0} p = \frac{4}{5}$.

We then define a function $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ from the two restrictions $(R - 1)$ and " $V_D = V_E$ " in $(R - 2)$.

$$F(\delta, q, \varepsilon) = \begin{pmatrix} f_1(\delta, q, \varepsilon) \\ f_2(\delta, q, \varepsilon) \end{pmatrix} = 0,$$

where

$$\begin{aligned} f_1(\delta, q, \varepsilon) &= \frac{4}{5}\delta(1 - 2\varepsilon)^2 V^{MM} + \delta \left[\frac{17}{5}\varepsilon - \frac{16}{5}\varepsilon^2 - \frac{4\varepsilon}{5p} \right] V^{MD} \\ &\quad + \frac{4\delta}{5}(\varepsilon - 4\varepsilon^2) V^{DM} - \frac{8}{5}(1 - \delta) \\ f_2(\delta, q, \varepsilon) &= \left[(1 - 2\varepsilon)^2 p - \varepsilon \right] V^{MM} - \left[\varepsilon - (1 - 2\varepsilon)^2 p \right] (V^{MD} + V^{DM}) + (1 - 3\varepsilon)p \end{aligned}$$

We can calculate³³ that

$$\begin{bmatrix} \frac{\partial f_1}{\partial \delta} & \frac{\partial f_1}{\partial q} \\ \frac{\partial f_2}{\partial \delta} & \frac{\partial f_2}{\partial q} \end{bmatrix}_{(\delta=\frac{10}{11}, q=\frac{1}{5}, \varepsilon=0)} = \begin{bmatrix} 4.476 & 2.365 \\ -0.507 & \frac{576}{6325} \end{bmatrix}.$$

³³We used Maple to calculate the numerical results. Also, note that evaluating q at $\frac{1}{5}$ is not exactly rigorous as we also have to make sure that $(1 - p - q) \in (0, 1)$ when $\varepsilon \in (0, \bar{\varepsilon})$. But we can always choose q to be arbitrarily close to $\frac{1}{5}$.

This matrix is invertible and our claim above is thus true by the implicit function theorem. In addition, we can further calculate that

$$\begin{aligned} \begin{pmatrix} \frac{d\tilde{\delta}}{d\varepsilon} \\ \frac{d\tilde{q}}{d\varepsilon} \end{pmatrix}_{(\delta=\frac{10}{11}, q=\frac{1}{5}, \varepsilon=0)} &= - \begin{bmatrix} \frac{\partial f_1}{\partial \delta} & \frac{\partial f_1}{\partial q} \\ \frac{\partial f_2}{\partial \delta} & \frac{\partial f_2}{\partial q} \end{bmatrix}^{-1} \begin{pmatrix} \frac{\partial f_1}{\partial \varepsilon} \\ \frac{\partial f_2}{\partial \varepsilon} \end{pmatrix}_{(\delta=\frac{10}{11}, q=\frac{1}{5}, \varepsilon=0)} \\ &= \begin{pmatrix} 0.767 \\ -0.4 \end{pmatrix}. \end{aligned}$$

This completes the verification of the claim above.

The last step is to prove **Proposition 3**. Before we formally prove the proposition, we point out the non-trivial fact that for any $\mu \in [0, \pi]$, we have that all the belief operators " $\chi^{Cc}(\mu)$, $\chi^{Cd}(\mu)$, $\chi^{Ce}(\mu)$, $\chi^{Dc}(\mu)$, $\chi^{Dd}(\mu)$, $\chi^{De}(\mu)$, $\chi^{Ec}(\mu)$, $\chi^{Ed}(\mu)$, $\chi^{Ee}(\mu)$ " have ranges as some subsets of $[0, \pi]$. For the belief operator $\chi^{Cc}(\mu)$, this is obvious. For all the other belief operators, note that all of them lie below the 45-degree line (for example, $\frac{d\chi^{Dc}(\mu)}{d\mu} > 0$, $\frac{d^2\chi^{Dc}(\mu)}{d\mu^2} < 0$). It is thus sufficient for us to only consider beliefs in $[0, \pi]$ for *any* player at *any* time.

Proof. We show that there is no profitable one-shot deviation for any player at any private information set if each player plays the full belief-based strategy (ρ, π) .

First, note that by construction, each player's belief is restricted in the interval $[0, \pi]$. We thus only consider two cases: $\mu = \pi$ and $\mu < \pi$.

Case 1 $\mu = \pi$.

Restriction $(R-1)$ ensures that a player's belief is fixed to be π in the class of histories (Cc, \dots, Cc) , and restriction $(R-2)$ ensures that any one-shot deviation is not profitable.

Case 2 $\mu < \pi$.

There are two possible one-shot deviations for this case: one is to play "C" in this period and continue with (ρ, π) , the other is to play "E" in this period and continue with (ρ, π) .

- The option of playing "C" now and then continuing with (ρ, π) :

Once the player has played "C", strategy (ρ, π) indicates that (1) the player play " σ^M " if the player sees signal "c" and $\chi^{Cc}(\mu) = \pi$, while play " σ^D " for all other cases; (2) the player play " σ^D " if the player sees signal "c" and $\chi^{Cc}(\mu) < \pi$, while play " σ^M " for all other cases. Now as (1) is equivalent to " σ^M ", and thus is not profitable since " σ^D " is always preferred to " σ^M " when $\mu < \pi$. For (2), the one-shot deviation thus only differs from " σ^M " at the information set "Cc". Denote the value difference of playing " σ^D " and playing this one-shot deviation to be $\Delta V(\mu)$ and we have ($\mu' = \chi^{Cc}(\mu)$):

$$\Delta V(\mu) = [V_{\sigma^D}(\mu) - V_{\sigma^M}(\mu)] - \delta [(1 - 2\varepsilon)\mu p + \varepsilon(1 - \mu p)] [V_{\sigma^D}(\mu') - V_{\sigma^M}(\mu')].$$

Now since $\Delta V(\pi) = 0$ and $\mu < \chi^{Cc}(\mu) < \pi$, it is immediate that $\Delta V(\mu) > 0$, or that the one-shot deviation is not profitable.

- The option of playing "E" now and then continuing with (ρ, π) :

Once the player has played "E", strategy (ρ, π) indicates that the player always play " σ^D ", no matter what signal she sees at the end of the current period. The value of playing "D" in this period and then continuing with (ρ, π) is:

$$V_D = \mu p \left(\frac{15}{4} (1 - \delta) + \delta \varepsilon V^{DM} \right) + (\mu q + 1 - \mu) (0) - \mu (1 - p - q) (1 - \delta).$$

And the value of playing "E" now and continuing with (ρ, π) is:

$$V_E = \mu p [3(1 - \delta) + \delta \varepsilon V^{DM}] - (\mu q + 1 - \mu)(1 - \delta) + \mu(1 - p - q)[-2(1 - \delta)].$$

Now given $\mu < \pi$, since

$$V_D - V_E = (1 - \delta) \left(1 - \frac{5}{4}\mu p\right) > 0 \text{ as } \mu < \pi \text{ and } \pi p = \frac{4}{5},$$

this one-shot deviation is not profitable either.

Hence, there is no profitable one-shot deviation for any player at any private information set. Together with the fact that the full belief-based strategy is realization equivalent to σ , we conclude that the full belief-based strategy profile $((\rho, \pi), (\rho, \pi))$ constitutes a sequential equilibrium. ■

References

- [1] D. ABREU, D. PEARCE, and E. STACCHETTI, "Optimal Cartel Equilibria with Imperfect Monitoring", *J. Econ. Theory* **39** (1986), 251-269.
- [2] D. ABREU, D. PEARCE, and E. STACCHETTI, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica* **58** (1990), 1041-1063.
- [3] I. AHN, "Imperfect information repeated games with a single perfect observer", *Mimeo*, 1996.
- [4] V. BHASKAR, "Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems", *Rev. of Econ. Stud.* **65** (1998), 135-149.
- [5] V. BHASKAR, "The Robustness of Repeated Game Equilibria to Incomplete Payoff Information", *Mimeo*, University of Essex (2000).
- [6] E. BEN-PORATH and M. KAHNEMAN, "Communication in Repeated Games with Private Monitoring", *J. Econ. Theory* **70** (1996), 281-297.
- [7] V. BHASKAR, G. MAILATH and S. MORRIS, "Purification in the Infinitely Repeated Prisoners' Dilemma ", *Mimeo*, (2006).
- [8] V. BHASKAR and I. OBARA, "Belief-based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring", *J. Econ. Theory* **102** (2002), 40-69.
- [9] V. BHASKAR and E. VAN DAMME, "Moral Hazard and Private Monitoring", *J. Econ. Theory* **102** (2002), 16-39.
- [10] O. COMPTE, "On sustaining cooperation without Public Observations", *J. Econ. Theory* **102** (2002), 106-150.
- [11] O. COMPTE, "On Failing to Cooperate when Monitoring is Private", *J. Econ. Theory* **102** (2002), 151-188.
- [12] G. ELLISON, "Cooperation in the Prisoners' Dilemma with Anonymous Random Matching", *Rev. Econ. Stud.* **61** (1994), 567-588.
- [13] J. ELY, J. HÖRNER and W. OLSZEWSKI, "Belief-Free Equilibria in Repeated Games", *Econometrica* **73** (2) (2005), 377-415.
- [14] J. ELY and J. VÄLIMÄKI, "A Robust Folk Theorem for the Prisoners' Dilemma", *J. Econ. Theory* **102** (2002), 84-105.
- [15] D. FUDENBERG, D. LEVINE and E. MASKIN, "The Folk Theorem with Imperfect Public Information", *Econometrica* **62** (1994), 997-1039.
- [16] D. FUDENBERG, and E. MASKIN, "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information", *Econometrica* **54** (1986), 533-556.
- [17] E. GREEN, and R. PORTER, "Noncooperative Collusion under Imperfect Price Information", *Econometrica*, **52** (1984), 87-100.

- [18] J. HARSANYI, "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points", *Int. J. Game Theory*, **2** (1973), 1-23.
- [19] J. HÖRNER and W. OLSZEWSKI, "The Folk Theorem for Games with Private Almost-Perfect Monitoring", *Econometrica*, forthcoming.
- [20] M. KANDORI, "Social Norms and Community Enforcement", *Rev. Econ. Stud.* **59** (1991), 63-80.
- [21] M. KANDORI, "Introduction to Repeated Games with Private Monitoring", *J. Econ. Theory* **102** (2002), 1-15.
- [22] M. PICCIONE, "The Repeated Prisoners' Dilemma with Imperfect Private Monitoring", *J. Econ. Theory* **102** (2002), 70-83.
- [23] G. MAILATH and S. MORRIS, "Repeated Games with Almost Public Monitoring", *J. Econ. Theory* **102** (2002), 189-228.
- [24] G. MAILATH and S. MORRIS, "Coordination Failure in Repeated Games with Almost Public Monitoring", *Theoretical Economics* **1** (2006), 311-340.
- [25] G. MAILATH AND L. SAMUELSON, "Repeated Games and Reputations: Long-Run Relationships". *Oxford University Press*, Oxford (2006).
- [26] H. MATSUSHIMA, "Repeated Games with Private Monitoring: Two Players", *Econometrica* **72** (3) (2004), 823-852.
- [27] M. KANDORI and I. OBARA, "Efficiency in Repeated Games Revisited: The Role of Private Strategies", *Econometrica* **74** (2) (2006), 499-519.
- [28] T. SEKIGUCHI, "Efficiency in Repeated Prisoners' Dilemma with Private Monitoring", *J. Econ. Theory* **76** (1997), 345-361.
- [29] T. SEKIGUCHI, "Robustness of Efficient Equilibria in Repeated Games with Imperfect Private Monitoring", *Mimeo*, (1998).
- [30] T. SEKIGUCHI, "Existence of Nontrivial Equilibria in Repeated Games with Imperfect Private Monitoring", *Games and Economic Behavior* (2002), 299-321.
- [31] G. STIGLER, "A Theory of Oligopoly", *J. of Political Econ.* **72** (1964), 44-61.
- [32] Y. YAMAMOTO, "Efficiency Results in N -Players Games with Imperfect Private Monitoring", *J. Econ. Theory* Forthcoming.
- [33] Y. YAMAMOTO, "A Characterization of Belief-Free Payoffs in Repeated Games with Almost Perfect Monitoring", *Mimeo* (2006).