Equilibrium Characterization of Repeated Games with Private Monitoring

Lei Gao *

March 17, 2018

Preliminary.

Abstract

This paper examines sequential equilibria of repeated games with private monitoring for very general distribution of private signals. Assuming full dimensionality conditions, we characterize the set of equilibrium values for 2-player games in which both actions and signals are finite. The folk theorem is partial and the equilibrium values are strictly bounded away from efficiency. The method is valid for N-player games, $N \geq 3$, but cumbersome in notation requiring high-dimensional array operations.

Keywords: Repeated game, prisoners' dilemma, private monitoring, sequential equilibrium.

JEL Classification: C72, C73, D82.

*School of Economics, Shandong University, Jinan, Shandong Province 250100, P. R. China; E-mail address: lgao@sdu.edu.cn.

1 Introduction

This paper examines sequential equilibria of repeated games with private monitoring, where actions and signals are finite, and the signals are generally distributed but under certain statistical identifiability conditions. We can construct a sequence of Bayesian stage games with consistently updated private beliefs from each strategy profile of a private monitoring game. The matrix representation of strategy profiles and strategy operators introduced in the construction assist us in describing the set of sequential strategies. We have fully characterized the set of equilibrium values for 2-player games and verified a version of folk theorem. The folk theorem is partial and the equilibrium payoffs are strictly bounded away from efficiency in general. Sufficiently informative private signals are necessary to characterize the equilibrium sets of N-player games and we need new techniques to investigate the strategy profiles requiring player-specific punishment.

In the characterization, a strategy profile leads to a sequence of Bayesian stage games with consecutively updated private beliefs. The payoffs of a player in a Bayesian stage game are her continuation promises when she believes that her opponents select each and every pure-action choice almost surely and she chooses a particular action. Beliefs are probabilities over the set of action choices of her opponents, and belief update at the end of a period and immediately before the next is consistent with players' plan of action choices after their histories according to the strategy. Value functions calculate expected continuation values given the continuation promises and beliefs. To solve for the value functions and continuation promises therein, we build a system of functional equations from a recovered recursive structure. A player's value functions, as functions of her beliefs, are solutions of these equations, and each of her actions identifies a value function.

We recover a recursive structure through an alternative view of the players' information sets. We introduce filtered space and filtration of histories so that players face continuation games strategically identical to the original game. Private beliefs of a player are adapted to her filtration and summarized by probability distributions over opponents' upcoming action choices. Belief update is built on strategy profiles and distribution of signals. Because a player's belief in a period is uniquely determined by her past outcome path once the initial belief and strategy are given, first-order beliefs determine all higher order beliefs. The first-order beliefs are sufficient statistics for the continuation of play.

A value function of a player in a period calculates the expectation of her continuation promises with first-order beliefs, so its functional form is a dot product of these two. Because of the recovered recursive structure, the value function can be decomposed into a weighted sum of the expected payoff from current play and values from future play, which are given by her value functions in the next period, functions of updated beliefs and a new set of continuation promises. A matrix representation of this decomposition gives us a clear view of the set of all possible strategy profiles. A strategy profile can be treated as a way players selecting continuation promises, so, in the space of value functions, strategies are operators acting on the functions in the decomposition. With the method of self-generation proposed by Abreu, Pearce, and Stacchetti (1990), this approach facilitates the characterization of equilibrium sets.

Apart from the trivial case of repeated play of stage game Nash equilibria, pure-strategy and mixed-strategy sequential equilibria behave differently. In a sequential pure-strategy profile, we show that each player, in each period, has a partition of her belief manifold, a subset of a Euclidean space collecting all her private beliefs. Over each submanifold, the pure-action choice assigned by the profile is the best response to the beliefs, or equivalently speaking, the value function identified by the chosen action is the maximal amongst all value functions. We say, these beliefs support that value function. And belief consistency means that updated supporting beliefs are supporting too. In terms of payoffs, if the profile sustains an inefficient stage game Nash equilibrium and a dominating action profile, the equilibrium payoffs converge to the Nash stage game payoffs when players become more and more patient. For a sequential mixed-strategy profile, beliefs are irrelevant in action choices, but they affect continuation values because the value functions associated with different action choices are identical but nontrivial as functions of beliefs. A version of folk theorem is obtained by taking convex hulls of the equilibrium value sets. Equilibrium values are strictly bounded away from efficiency in general.

This paper provides a feasible implementation of Amarante (2003)'s idea that private monitoring games admit a recursive structure, and builds its connection with the work by Phelan and Skrzypacz

(2012), and Kandori and Obara (2010). These papers focus on finite-state strategies. From the view of filtered spaces, players' belief manifolds and submanifolds are embedded in Euclidean spaces, so it is sufficient to check incentives only for extreme beliefs. Moreover, the belief sequences are contained in compact manifolds, which guarantees convergent subsequences. The limit beliefs are the fixed points of the set-based operators in Phelan and Skrzypacz (2012). Our method is also an alternative to the one proposed by Iwasaki, Joe, Kandori, Obara, and Yokoo (2012).

As discussed in Kandori (2002), private monitoring games have been explored from many perspectives in the literature. Compte (1998) and Kandori and Matsushima (1998) introduce public communication to recover a recursive structure in private monitoring games, while Bhaskar and Obara (2002) and Mailath and Morris (2002, 2006) consider almost public monitoring, where there are small private perturbations in public monitoring. Mailath and Morris (2002) investigate the robustness of perfect public equilibria (PPE) to the introduction of small degrees of private monitoring. They show that the strategies with bounded recall are robust to perturbations, whereas infinite-history dependent strategies are not. Sekiguchi (1997) provides the first example of efficiency in private monitoring repeated games where the monitoring is sufficiently close to perfect. Piccione (2002) introduces belief-free equilibria, which have been the focus of extensive research, e.g., Ely and Välimäki (2002), Ely, Hörner, and Olszewski (2005), and Kandori and Obara (2006).

The paper is organized as follows. Section 2 formally presents the model and constructs filtered space, filtration, beliefs, belief operators, and value functions. This section also gives a private-monitoring version of the one-shot deviation principle. Section 3 provides a practical approach to verify intertemporal incentives of a strategy profile. Section 4 examines sequential equilibrium profiles of the repeated prisoners' dilemma under private monitoring. Section 5 follows the method of self-generation by Abreu, Pearce, and Stacchetti (1990) to characterize the value function space and set of equilibrium values, $\mathcal{E}(\delta)$. Section 6 verifies the folk theorem. Section 7 concludes. The appendix lists related notations and provides major proofs.

2 Private Monitoring Games: The Model

We analyze infinitely repeated games with imperfect private monitoring, which are infinite repetition of a stage game where the observation of opponents' actions is noisy and private. In this section we will present the model and lay out a theoretical foundation, and in the following sections we will introduce matrix representation of strategy profiles and strategy operators through the example of repeated prisoners' dilemma. It facilitates our characterization of equilibrium sets.

There are n long-lived players, player i, $i \in \{1, 2, \ldots, n\}$, and each player has a finite stage game action set A_i. At the end of each period, player i observes a private signal z_i, drawn from a finite set Z_i. The probability distribution of the private signal vector $z = (z_1, z_2, \ldots, z_n) \in Z \equiv \prod_{i=1}^{n} Z_i$ is given by $\pi(z|a)$, where a is the action profile in $A \equiv \prod_{i=1}^{n} A_i$. The marginal distributions of private signals are given by $\pi_i(\cdot|a)$. The private signal z_i is the only information player i can receive about the behavior of other players at each stage game. Throughout this paper we assume that all players use the same discount factor, δ, but we do not assume full support of the signals unless otherwise stated.

Following the literature, we assume that the stage game ex-post payoff of player i after the realization (z, a) is given by $u_i^*(z_i, a_i)$. The ex-ante stage game payoffs are given by $u_i(a) \equiv \sum_z u_i^*(z_i, a_i)\pi_i(z_i|a)$. Then a repeated game is indexed by the triplet (Z, π, u) for a given set of players and their action choices A.

Here are the definitions that we need in analyzing private monitoring games. For intuition behind these definitions, please refer to the two-period example in Appendix B.

Definition 1. *A state space Ω_i for player i is a set which collects all possible outcomes of the infinitely repeated game observed by player i, $\Omega_i \equiv \{\omega_i \mid \omega_i = a_i^0 z_i^0 a_i^1 z_i^1 \ldots a_i^t z_i^t \ldots\}$, where $a_i^t \in A_i$, $z_i^t \in Z_i$, $t = 0, 1, 2, \ldots, \infty$. The element, ω_i, is an infinite series of player i's actions and signals formed in chronological order, by which player i receives information as the game proceeds. Let $\Omega \equiv \prod_{i=1}^{n} \Omega_i$.*

The imperfect public monitoring games are only a special case of the private monitoring games where $z_i = z_j$, for all $a \in A$ and all t, that is, the signals observed by all players are always the same. When the monitoring is truly private, no player can observe a generic element in Ω.

Definition 2. *A private history h_i^t, $t = 1, 2, \ldots, \infty$, for player i is a subset of player i's state space Ω_i. A given history h_i^t, $h_i^t \subseteq \Omega_i$, contains all of the elements that have the same first $2t$ entries representing the actions player i has chosen and her private signals for period $0, \ldots, t-1$. A history $h^t \subseteq \Omega$ is $\prod_{i=1}^n h_i^t$.*

In the game of repeated prisoners' dilemma, for example,

$$h_i^2 = \{\omega \in \Omega_i \mid \omega = a_i^0 z_i^0 a_i^1 z_i^1 a_i^2 z_i^2 \ldots,\ a_i^0 = E, z_i^0 = \bar{z}, a_i^1 = E, z_i^1 = \underline{z}\}$$

is a history in which player i plays E, E and observes $\bar{z}$, $\underline{z}$ for periods 0 and 1. The histories h_i^t, or h^t, are the realized outcome *paths* from period 0 to period $t-1$.

Let $\mathcal{F}_i^0 \equiv \{\phi, \Omega_i\}$ denote the trivial σ-algebra defined on Ω_i, and define $\mathcal{F}_i^t$, $t \geq 1$, to be the σ-algebra on Ω_i and generated by h_i^t.

Definition 3. *A filtration $\mathfrak{F}_i$ is a sequence of σ-algebras, $\{\mathfrak{F}_i^t\}_{t=0}^\infty$, $\mathfrak{F}_i^t \equiv \prod_{j=1}^n \mathcal{F}_j^\tau$, $\tau = t$ for $j = i$ and 0 otherwise, that is, $\mathfrak{F}_i^t \equiv \mathcal{F}_1^0 \times \ldots \times \mathcal{F}_i^t \times \ldots \times \mathcal{F}_n^0$. Let $\mathfrak{F} = \{\mathfrak{F}_i\}_{i=1}^n$.*

For $s < t$, the fact that $\mathcal{F}_i^s \subset \mathcal{F}_i^t$ and $\mathfrak{F}_i^t$ is a finer σ-algebra than $\mathfrak{F}_i^s$ indicates that player i has more information in period t than in period s. Note that, $\mathfrak{F}_i^t = \mathcal{F}_1^0 \times \ldots \times \mathcal{F}_i^t \times \ldots \times \mathcal{F}_n^0$, $t = 0, 1, 2, \ldots, \infty$, characterizes the private monitoring: Player i knows her history up to period t, but she is uncertain about player j's history.

Definition 4. *A behavior strategy for player i is a sequence of functions $\mathfrak{s}_i \equiv \{\mathfrak{s}_i^t\}_{t=0}^\infty$, $\mathfrak{s}_i^t : \Omega \rightarrow \Delta(A_i)$, which are* adapted *to the filtration $\mathfrak{F}_i$, that is, $\mathfrak{s}_i^t$ is measurable with respect to $\mathfrak{F}_i^t$.*

In general, for $s < t$, $\mathfrak{s}_i^t$ may not be measurable with respect to $\mathfrak{F}_i^s$, which means $\mathfrak{s}_i^t$ is only a plan for player i in period s. When restricted to Ω_j, $j \neq i$, $\mathfrak{s}_i^t$ for all t is solely measurable with respect to $\mathcal{F}_j^0$, the trivial σ-algebra defined on Ω_j, so strategy $\mathfrak{s}_i$ is private. The strategy profile $\mathfrak{s} \equiv \{\mathfrak{s}_i\}_{i=1}^n$ may *not* be action free as discussed in Mailath and Samuelson (2006).

Definition 5. *The beliefs of player i are a sequence of probability distributions $\mathfrak{B}_i \equiv \{\mathfrak{B}_i^t\}_{t=0}^\infty$ over Ω.*

At the beginning of period 0, player i holds an initial belief (probability distribution) $\mathfrak{B}_i^0$ of the elements in Ω. For $t = 1, 2, \ldots, \infty$, $\mathfrak{B}_i^t$ assigns probability 1 to some history h_i^t, while $\mathfrak{B}_i^t|_{\prod_{j \neq i} \Omega_j}$ is player i's updated belief over h_j^t. Because the number of outcome paths in Ω is no less than the set of real numbers, precisely writing down the probability distributions $\mathfrak{B}$ will lead to unnecessary complication. This motivates the following definition.

Definition 6. *The belief $\mathfrak{B}_i$ is progressively adapted to the filtration $\mathfrak{F}_i$ if $\mathfrak{B}_i^t$ is a discrete probability distribution over h^t for all t.*

In summary, the σ-algebra $\mathfrak{F}_i^t$ is a product of $\mathcal{F}_i^t$ and $\mathcal{F}_j^0$, $j \neq i$, and those h_i^t are generating sets of $\mathcal{F}_i^t$. Player i observes h_i^t but is uncertain about h_j^t. The progressive adaptiveness is necessary and sufficient to characterize the state of uncertainty. More specifically, progressively adapted $\mathfrak{B}_i^t$ assigns probability 1 to some h_i^t, while $\mathfrak{B}_i^t|_{\prod_{j \neq i} \Omega_j}$ is a discrete probability distribution over h_j^t.

We define second-order beliefs next.

Definition 7. *The beliefs of beliefs for player i are a sequence of mappings $\tilde{\mathfrak{B}}_i \equiv \{\tilde{\mathfrak{B}}_i^t\}_{t=0}^{\infty}$, where $\tilde{\mathfrak{B}}_i^t$ maps each h_i^t to a probability distribution over player j's beliefs $\mathfrak{B}_j^t$, $j = 1, \ldots, n$, and $j \neq i$.*

Because of the special structure of player beliefs, it is easy to verify the following connection between belief and belief of beliefs.

Lemma 1. *For a given h_i^t, $i \in \{1, 2, \ldots, n\}$ and $t \geq 1$, the probability distribution $\tilde{\mathfrak{B}}_i^t(h_i^t)$ coincides with $\mathfrak{B}_i^t|_{\prod_{j \neq i} \Omega_j}$, where $\mathfrak{B}_i^t$ assigns probability 1 to the given h_i^t.*

Lemma 1 establishes that the belief system $\mathfrak{B} \equiv \{\mathfrak{B}_i\}_{i=1}^n$ determines all higher order beliefs for all players. More importantly, different from public monitoring games in which observation of other players' behavior is common, in private monitoring games, it is the formation of the belief system, $\mathfrak{B}$, that is common knowledge.

For notational simplicity, let β or β_i denote player i's belief, the probability distribution over h_j^t; let β_i^t specifically denote $\mathfrak{B}_i^t|_{\prod_{j \neq i} \Omega_j}$, the initial belief in stage t, i, $j \in \{1, 2, \ldots, n\}$, $j \neq i$, and $t = 0, 1, \ldots, \infty$; and let $\beta(h_i^t)$ be the belief representing the image of the map $\tilde{\mathfrak{B}}_i^t$ for a given history

h_i^t. Note that it is the strategy profile $\mathfrak{s}$ that determines β and, given t, $\beta(h_i^t)$ is different for different h_i^t in general. We record the pair $(\mathfrak{s}_i^t, \beta_i^t)$ explicitly in case we need to show their connection.

Define statistical inference operation $\mathcal{I}(h, \beta)$.

Definition 8. *A statistical inference $\mathcal{I}(h_i^t, \beta_i)$ by player i is her updated belief about her opponent's private histories, given initial belief β_i and history h_i^t, $i \in \{1, 2, \dots, n\}$ and $t \geq 1$.*

In the definition, h_i^t could be any segment of an outcome path not necessarily starting from the initial period, only β_i must be the belief held by player i immediately before h_i^t is started.

We now define history concatenation. Because h_i^t is setwise identical to Ω_i for any $i \in \{1, 2, \dots, n\}$ and all t, a history $h_i^{t'}$, $t' \geq 1$, can be viewed either as a subset of Ω_i or a subset of h_i^t. As a subset of Ω_i, $h_i^{t'}$ is a history played from the very beginning of the repeated game, while as a subset of h_i^t, it is a history played immediately after the history h_i^t. The definition of history concatenation is given as follows.

Definition 9. *A concatenation operation $\circ$ between two histories h_i^t and $h_i^{t'}$, $h_i^t \circ h_i^{t'}$, is a joint operation of the two histories, and is a subset of Ω_i with the first $2t$ entries of its elements identical to those of h_i^t and $2t + 1$ to $2t + 2t'$ entries of the elements identical to the first $2t'$ entries of $h_i^{t'}$.*

It is easy to show the following result about statistical inference over concatenated histories.

Lemma 2. $\mathcal{I}(h_i^t \circ h_i^{t'}, \beta_i) = \mathcal{I}\big(h_i^{t'}, \mathcal{I}(h_i^t, \beta_i)\big)$. *If we write history h_i^t as one-stage history concatenation $h_i^{(0)} \circ h_i^{(1)} \circ \dots \circ h_i^{(t-1)}$, then $\mathcal{I}(h_i^t, \beta_i) = \mathcal{I}\big(h_i^{(t-1)}, \dots \mathcal{I}(h_i^{(0)}, \beta_i) \dots\big)$.*

Proof. A player's private signals are serially uncorrelated, so statistical inference after any two non-overlapping histories is independent. The claim follows. □

The statistical property of player private histories is summarized by the following proposition.

Proposition 1. *Given the conditional signal distributions $\pi(\cdot|a)$ for any action profile $a \in A$, the history h_i^t for player i, $i \in \{1, 2, \dots, n\}$ and $t \geq 1$, is a sufficient statistic for player i's beliefs $\mathfrak{B}_i^t$, $t \geq 1$, and they are inferred by $\mathcal{I}(h_i^t, \mathfrak{B}_i^0)$.*

We now construct the recursive structure of strategy profiles in repeated games with private monitoring. The automaton representation of the profiles here has one more ingredient than those of perfect monitoring or imperfect public monitoring games, namely, the evolving process of beliefs β_i^t.

We start from the standard automaton representation of public monitoring games. We need a separate automaton for each player because observation is private. A behavior strategy $\mathfrak{s}_i$ is represented by a set of states $\mathcal{W}_i$, in which the states are collections of histories h_i^t, $t = 1, 2, \ldots, \infty$, an initial state w_i^0, a decision rule $f_i : \mathcal{W}_i \to \Delta(A_i)$ specifying a distribution over action choices for each state, and a transition function $\tau_i : \mathcal{W}_i \times A_i \times Z_i \to \Delta(\mathcal{W}_i)$. Because the profiles are private and could be mixed strategies, the transitions to the next period states are potentially random. A state w' is *accessible* from another state w if there is some history path h_i^t such that $w' = \tau_i(w, h_i^t)$, where $\tau_i(w, h_i^t)$ is defined recursively as $\tau_i(w, h_i^t) = \tau_i(\tau_i(w, h_i^{t-1}), a_i^{t-1}, z_i^{t-1})$.

There is a separate belief process $\{\beta_i^t\}_{t=0}^{\infty}$ for each player. The initial belief β_i^0 is *not* required to be consistent with w_j^0, her opponents' initial states, because this type of consistency cannot be maintained beyond the first period. The distributions $\{\beta_i^t\}_{t=0}^{\infty}$ is given by $\beta_i^{t+1} = \mathcal{I}(h_i^{(t)}, \beta_i^t)$, where $h_i^{(t)} = a_i^t z_i^t$ is player i's period t action choice and observed private signal. We treat $\mathcal{I}(a_i^t z_i^t, \cdot)$ as belief operators $a_i z_i^*$, where $a_i \in A_i$, and $z_i \in Z_i$.

Definition 10. *The belief operator $a_i z_i^*$, where $a_i \in A_i$, and $z_i \in Z_i$, is a Bayesian information updating process,*

$$\beta_i^{t+1} = a_i z_i^*(\beta_i^t), \quad t = 0, 1, \ldots, \infty.$$

β_i^t is player i's initial belief of player j's states at the beginning of period t; and $\beta_i^{t+1} = a_i z_i^*(\beta_i^t)$ is based on player i's action choice a_i and the following received signal z_i, player i's updated belief of player j's would-be states at the end of period t (also the beginning of period $t+1$).

The updated belief β_i^{t+1} at the end of period t is player i's initial belief for period $t+1$. In the inference, a_i is the output of the decision rule f_i when player i observes her state at the beginning of period t, while player j chooses a_j following f_j with probabilities assigned by β_i^t from player i's standpoint. Because the monitoring is private, player i relies on $a_i z_i$ to update her beliefs over $a_j z_j$.

The exact procedure is as follows, player i updates her beliefs on the choices of a_j based on a_i and z_i, then derives a distribution of z_j conditioned on a_j's updated distribution and observed z_i; after that, β_i^{t+1} is given by τ_j. So it is a combination of f_i, f_j, τ_j, conditional probabilities π, and beliefs β_i^t over $\prod_{j \neq i} \mathcal{W}_j$ that gives β_i^{t+1}.

Although beliefs $\{\beta_i^t\}_{t=0}^{\infty}$ are important to a strategy profile, they are not an independent component of its automaton representation $(\mathcal{W}_i, w_i^0, f_i, \tau_i)$, $i = 1, 2, \ldots, n$. A strategy profile may assign purely belief-dependent action choices, but it can still be written in a standard automaton representation because once the initial beliefs are given, any finite-period history of a player uniquely identifies a private belief of hers.

Proposition 2. *Any purely belief-dependent strategy profile can be expressed by a standard automaton representation.*

Proof. See Appendix. □

Finally, by Zorn's lemma, we can always find the minimal automata, and the minimal automaton for each player has to be unique. Other automaton representations will result in identical value functions for different states. We only consider minimal automaton representations from now on.

We define $V_i(w_i, \beta_i)$ for player i, a function of state w_i and belief β_i, as the expected discounted continuation value from play that begins in state w_i and with belief β_i, if all players follow decision rules f. This function is not dependent on i's opponent j's state w_j, belief β_j, and action choice a_j, $j \neq i$, because all of the information player i has about player j is based on β_i, which is also player i's belief of player j's state-dependent beliefs by Lemma 1. The bottom line is that player i knows that player j will follow the decision rule f_j, for any j, $j \neq i$.

Definition 11. *Suppose that the strategy profile $\mathfrak{s}$ is described by automata $(\mathcal{W}_i, w_i^0, f_i, \tau_i)$, $i = 1, 2, \ldots, n$. Given player i's beginning state $w_i = w$ and belief $\beta_i = \beta$, the value of continuation*

$V_i(w,\beta)$ *for player* i *when all players follow the decision rules* f_j*,* $j=1,2,\dots,n$*, is*

$$\begin{aligned} V_i(w,\beta) &= (1-\delta)\mathbb{E}^{\beta}_{f_i,f_j}\big[u_i(a)\big] + \delta\ \mathbb{E}^{\beta,\pi}_{f_i,f_j}\Big[V_i\big(\tau_i(w,a_i,z_i),a_i z_i^*(\beta)\big)\Big] \\ &= (1-\delta)\sum_{a_i,a_j} u_i(a) Prob\big(a=a_i a_j|f_i;\beta,f_j\big) \\ &\quad + \delta \sum_{a_i,z_i} V_i\big(\tau_i(w,a_i,z_i),a_i z_i^*(\beta)\big) Prob\big(a_i z_i|f_i;\beta,f_j;\pi_i(z_i|a)\big). \end{aligned} \tag{2.1}$$

When the monitoring has full support, every history arises with positive probability. It is straightforward to derive the incentive condition for sequentially rational equilibria under private monitoring. The following proposition is, in fact, a rephrase of the original one-shot deviation principle.

Proposition 3. *(The one-shot deviation principle) A strategy profile* $\mathfrak{s}$ *is sequentially rational if and only if there are no profitable one-short deviations, that is, if and only if*

$$\begin{aligned} V_i(w,\beta) \geq\ & (1-\delta)\sum_{a_j} u_i(a) Prob\big(a=a_i a_j|\beta,f_j\big) \\ & + \delta \sum_{z_i} V_i\big(\tau_i(w,a_i,z_i),a_i z_i^*(\beta)\big) Prob\big(a_i z_i|\beta,f_j;\pi_i(z_i|a)\big), \quad i=1,2,\dots,n, \end{aligned}$$

for any $a_i \in A_i$ *and any possible* β *that can be reached when player* i *is in state* w*.*

The deviation is one-shot because it occurs only in the current period and from the next period on the players revert to the decision rules.

3 Verification of Intertemporal Incentives

In this section, we mainly examine the value functions and private beliefs of pure-strategy equilibrium profiles. The behaviors of mixed-strategies will become clear once the work about pure-strategies is done. Their properties will assist us in verifying sequential rationality. We follow the classical method involving automaton representations of strategy profiles at this stage. After the example of repeated prisoners' dilemma in the next section, we will introduce matrix representations and rephrase our results in the language of operators.

In the automaton representation of a pure-strategy profile, the range of a player's private beliefs forms a manifold in a Euclidean space, while her value functions are hyperplanes defined over the manifold and identified by her states. For the strategy profile to be sequentially rational, each player has a partition of her belief manifold, and over each submanifold, a hyperplane associated with certain state is maximal amongst all hyperplanes. Equivalently speaking, the states of a player divide her private beliefs into regions, and over each region the action choices assigned to the states by the strategy profile are the best response to the beliefs in that region.

The description is valid for strategies that have automaton representations with countably infinite number of states, because both action and signals are finite – there are only finitely many possibilities to be considered in each period. It will not be an issue for mixed-strategies as to be shown, because their value functions are identical in any period for a given strategy profile and they can have uncountably many states.

There are two types of consistency necessary for sequential rationality. When a player chooses an action for a state in a strategy profile, she is expecting mixed actions from her opponents because she is uncertain about their states under imperfect private monitoring. Because beliefs are over each other's states and states are assigned with action choices, consistency requires that beliefs of a player combined with their strategy profile yield probabilities over actions. A second consistency is between states and beliefs in continuation values, which we will explain in details below.

Although value function $V_i(w, \beta)$ is a function of belief β well-defined over any possible value of beliefs, for $V_i(w, \beta)$ to represent a continuation value, the state w and belief β must be consistent. That is to say that there exists a history h_i^t such that $h_i^t \in w$ and $\mathfrak{B}_i^t|_{\prod_{j\neq i}\Omega_j} = \beta$ for the given h_i^t. We define that β is *reachable* when player i is in state w. Each state of a player determines a set of reachable beliefs of her opponent's states. Let β^{w_i} denote the set of player i's beliefs consistent with her state w_i, the reachable beliefs for some i's history in state w_i.

We characterize the sets $\{\beta^w\}_{w\in\mathcal{W}_i}$ of player i, $i = 1, 2, \ldots, n$, for sequentially rational pure-strategies. We temporarily consider finite-state automaton representations at this point. This won't affect the characterization because only finitely many states occur in each period. A belief β_i^t can be written as a vector $(\beta_i^{t,(1)}, \dots, \beta_i^{t,(\eta_i)})$, which denotes a discrete probability distribution over the finite η_i state-vectors of player i's opponents, $|\prod_{j\neq i} \mathcal{W}_j| = \eta_i$. The restrictions

$$\beta_i^{t,(1)} + \ldots + \beta_i^{t,(\eta_i)} = 1, \quad \beta_i^{t,(k)} \geq 0, \quad k = 1, \dots, \eta_i,$$

define $(\eta_i - 1)$-dimensional player i's manifold $\mathcal{M}_i$ in an η_i-dimensional Euclidean space $\mathbb{R}^{\eta_i}$. This manifold is exactly the range of the belief $\mathfrak{B}_i^t|_{\prod_{j\neq i}\Omega_j}$. We introduce a convex and connected partition $\mathcal{P}_i$ of $\mathcal{M}_i$.

Definition 12. *A convex and connected partition $\mathcal{P}_i$ of the belief manifold $\mathcal{M}_i$ is a collection of convex and connected submanifolds $\mathcal{M}_i^{(1)}, \dots, \mathcal{M}_i^{(\kappa_i)}$, where κ_i is the number of states in $\mathcal{W}_i$, $|\mathcal{W}_i| = \kappa_i$, such that $\bigcup_{k=1}^{\kappa_i} \mathcal{M}_i^{(k)} = \mathcal{M}_i$ and any pairwise intersection of $\mathcal{M}_i^{(k)}$, $k = 1, 2, \dots, \kappa_i$, is empty.*

Here is a property that the sets $\{\beta^w\}$ should have when the corresponding pure-strategy profile is sequentially rational.

Proposition 4. *If a pure-strategy profile that has a finite-state automaton representation is sequentially rational, then for each player i, $i = 1, 2, \dots, n$, there exists a convex and connected partition $\mathcal{P}_i$ of her belief manifold $\mathcal{M}_i$, such that for each state $w \in \mathcal{W}_i$, the set β^w is contained in one of the submanifolds of $\mathcal{P}_i$ and no two such sets are contained in one submanifold.*

Without loss of generality, we re-index the submanifolds so that $\mathcal{M}_i^{(k)}$ contains $\beta^{w^{(k)}}$, $w^{(k)} \in \mathcal{W}_i$, $k = 1, \dots, \kappa_i$.

En route to proving Proposition 4, we show a representation of the continuation value, $V(w, \beta)$.

Proposition 5. *In a finite-state automaton representation of a strategy profile, for each $w_i \in \mathcal{W}_i$, player i has a corresponding function v_i which assigns a value $v_i(\prod_{j\neq i} w_j)$ to each state vector $\prod_{j\neq i} w_j \in \prod_{j\neq i} \mathcal{W}_j$, such that*

$$V_i(w_i, \beta) = \mathbb{E}^{\beta}\left[v_i\right] = \beta \cdot v_i = \sum_{k=1}^{\eta_i} \beta^{(k)} v_i(\prod_{j\neq i}^{(k)} w_j),$$

where $\beta = (\beta^{(1)}, \ldots, \beta^{(\eta_i)})$ *is the belief vector over* $\prod_{j \neq i} \mathcal{W}_j$*,* $|\prod_{j \neq i} \mathcal{W}_j| = \eta_i$*, and* $\prod_{j \neq i}^{(k)} w_j$ *goes through all elements in* $\prod_{j \neq i} \mathcal{W}_j$*.*

Please see the Appendix for the proofs of Proposition 5 and Proposition 4.

Remark. Given states w_i, w_j of player i and j, the value $v_i(\prod_{j \neq i} w_j)$ is an expected continuation value if player i is in state w_i and believes that her opponents are in state $\prod_{j \neq i} w_j$ almost surely. Proposition 5 is true for both pure and mixed strategies. Here, we only prove it for pure-strategies.

The linearity of value function $V_i(w, \beta)$ in belief β is implied by equation (2.1) in Definition 11. In the second term on the right hand side of equation (2.1), the portion of continuation values related to $a_i z_i$ depends on the joint distribution of $a_i z_i$ and the states of player i's opponents. Under sequential rationality, the linearity of continuation value imposes restrictions on β^w, which is characterized by Proposition 4. Sequential mixed-strategies behave differently but the characterization is straightforward.

In sequential mixed-strategy profiles, as functions of private beliefs, the value functions of different action choices assigned to the same state have identical functional form. So are those of the states after transition from the same state in previous period. The identical value functions guarantee that players are indifferent in action choices and state transitions regardless of their beliefs, but the functions themselves are belief-dependent. This resembles the *belief-free* equilibrium strategy profiles in the literature.

To summarize, in the minimal automaton representation of a pure-strategy profile, the value functions of a player identified by her states are hyperplanes of a Euclidean space defined over the manifold of her private beliefs. The strategy profile is sequentially rational if and only if each player has a partition of her belief manifold, and over each submanifold, a hyperplane which belongs to a certain state is maximal amongst all hyperplanes, that is, for each player, the action choice assigned to a state is the best response to her beliefs reachable from that state. In sequential mixed-strategy profiles, the value functions of a player in any period reduce to a unique form and partition of beliefs is no longer relevant.

4 Repeated Prisoners' Dilemma with Private Monitoring

We use the repeated prisoners' dilemma with imperfect private monitoring as an example to illustrate our construction that, in private monitoring games, players are playing Bayesian stage games in each period with consistently updated beliefs, and payoffs in the games are continuation promises which may vary in different periods.

Section 4.1 and 4.2 examine two classes of pure-strategy profiles whose counterparts are studied in perfect monitoring and imperfect public monitoring games. Section 4.3 studies a mixed-strategy profile. Section 4.4 discusses the matrix representation of strategy profiles and strategy operators acting on value functions. For symmetric strategies, we suppress the subscript i of players.

4.1 Grim Trigger

In the example of repeated prisoners' dilemma, the ex ante payoffs of a stage game are given in Figure 1. We fix ex ante payoffs to simplify calculations.

Figure 1: The prisoners' dilemma

	E	S
E	2,2	-1,3
S	3,-1	0,0

The ex ante payoffs of a stage game in the repeated prisoners' dilemma.

The monitoring is private. Player i, $i = 1, 2$, observes a *private* signal z_i drawn from the space $\{\underline{z}, \bar{z}\}$ in each stage game. The probability distribution of the private signal vector $(z_1, z_2) = \{\underline{z}, \bar{z}\}^2$ is given by $\pi(z_1 z_2|a)$, where $a \in A$ is an action profile. Let $\bar{z}_i$ denote $z_i = \bar{z}$, and similarly for $\underline{z}_i$. For simplicity, we assume symmetry in the marginal distribution of private signals, and it is given by (4.1),

$$\pi(\bar{z}|a) = \begin{cases} p, & \text{if } a = EE, \\ q, & \text{if } a = ES \text{ or } SE, \\ r, & \text{if } a = SS, \end{cases} \tag{4.1}$$

where $0 < q < p < 1$ and $0 < r < p$. In terms of the joint distribution of private signals, we combine four equations (4.2),

$$\begin{cases} \pi(\bar{z}_1\bar{z}_2|a) + \pi(\bar{z}_1\underline{z}_2|a) = \pi(\bar{z}_1|a), \\ \pi(\underline{z}_1\bar{z}_2|a) + \pi(\underline{z}_1\underline{z}_2|a) = \pi(\underline{z}_1|a), \\ \pi(\bar{z}_1\bar{z}_2|a) + \pi(\underline{z}_1\bar{z}_2|a) = \pi(\bar{z}_2|a), \\ \pi(\bar{z}_1\underline{z}_2|a) + \pi(\underline{z}_1\underline{z}_2|a) = \pi(\underline{z}_2|a). \end{cases} \tag{4.2}$$

The equations are under-specified to solve for the distribution, so we introduce a correlation coefficient ρ. The use of signal parameter ρ is for ease of exposition, and it is not necessary for equilibrium characterization of repeated games in general. Consider the conditional probability $\pi(\bar{z}_1|\bar{z}_2, a)$. We know $\pi^2(\bar{z}|a) \leq \pi(\bar{z}_1\bar{z}_2|a) \leq \pi(\bar{z}|a)$, while $\pi(\bar{z}_1\bar{z}_2|a) = \pi^2(\bar{z}|a) = \pi(\bar{z}_1|a)\pi(\bar{z}_2|a)$ indicates *independence* of *good* signals, and $\pi(\bar{z}_1\bar{z}_2|a) = \pi(\bar{z}|a)$, which implies $\pi(\bar{z}_1|\bar{z}_2, a) = \pi(\bar{z}_1\bar{z}_2|a)/\pi(\bar{z}_2|a) = 1 = \pi(\bar{z}_2|\bar{z}_1, a)$, indicates *perfectly correlation* of *good* signals. Define

$$\rho = \frac{\pi(\bar{z}_1\bar{z}_2|a) - \pi^2(\bar{z}|a)}{\pi(\bar{z}|a) - \pi^2(\bar{z}|a)}. \tag{4.3}$$

We have $\rho \in [0, 1]$ in which $\rho = 0$ and $\rho = 1$ represent the above two cases, respectively. Solve (4.2) and the joint distribution of the private signals is given by (4.4)

$$\begin{aligned} \pi(\bar{z}_1\bar{z}_2|a) &= \rho\pi(\bar{z}|a) + (1-\rho)\pi^2(\bar{z}|a), \\ \pi(\bar{z}_1\underline{z}_2|a) &= \pi(\underline{z}_1\bar{z}_2|a) = (1-\rho)(1-\pi(\bar{z}|a))\pi(\bar{z}|a), \\ \pi(\underline{z}_1\underline{z}_2|a) &= (1-\pi(\bar{z}|a)) - (1-\rho)(1-\pi(\bar{z}|a))\pi(\bar{z}|a). \end{aligned} \tag{4.4}$$

The conditional probabilities for private signals are:

$$\begin{aligned}
\pi(\bar{z}_i|a,\bar{z}_j) &= 1-(1-\rho)\big(1-\pi(\bar{z}|a)\big), \\
\pi(\underline{z}_i|a,\bar{z}_j) &= (1-\rho)\big(1-\pi(\bar{z}|a)\big), \\
\pi(\bar{z}_i|a,\underline{z}_j) &= (1-\rho)\pi(\bar{z}|a), \\
\pi(\underline{z}_i|a,\underline{z}_j) &= 1-(1-\rho)\pi(\bar{z}|a),
\end{aligned} \tag{4.5}$$

where $a = EE$, SE (or ES), SS, and i, $j \in \{1, 2\}$, $i \neq j$.

We now look at its sequential equilibria. Because mutual shirking is a stage-game Nash equilibrium, both players can simply ignore private signals and shirk in every period. To create intertemporal incentives for the players to exert effort, we need to associate low continuation values with certain action choices and realization of signals.

One of the simplest profiles is the *grim trigger* with the name borrowed from public monitoring games. The profile is symmetric and requires that both players exert effort in the first period and continue to exert effort only when the player has exerted effort and received a good signal $\bar{z}$. The strategy has a simple two-state automaton representation.[1] The state space is $\mathcal{W} = \{w^R, w^P\}$, where w^R denotes the state of *rewarding* and w^P denotes the state of *punishment*. The initial state is w^R, with the output function

$$f(w^R) = E, \quad \text{and} \quad f(w^P) = S,$$

and transition function

$$\tau(w, a_i z_i) = \begin{cases} w^R, & \text{if } w = w^R, \text{ and } a_i z_i = E\bar{z}, \\ w^P, & \text{otherwise,} \end{cases}$$

where a_i, z_i are the player's current period action and signal, respectively.

[1] For the construction of the filtration and filtered space, please refer to Section B in the Appendix for a similar procedure.

Let β_i^t denote the probability that player i's belief $\mathfrak{B}_i^t$ assigns to the case of player j being in rewarding state w^R when player i experiences h_i^t. The operators $a_i z_i^*$, $\beta_i^{t+1} = a_i z_i^*(\beta_i^t)$, where $a_i = E$, S, $z_i = \bar{z}$, $\underline{z}$, $i = 1$, 2, $t = 0, 1, 2, \ldots$, are functions acting on probability β. By Bayes' rule and (4.5), the functions are given as

$$\begin{aligned} E\bar{z}^*(\beta) &= \frac{\beta p}{\beta p + (1-\beta)q}(1-(1-\rho)(1-p)), \\ E\underline{z}^*(\beta) &= \frac{\beta(1-p)}{\beta(1-p)+(1-\beta)(1-q)}(1-\rho)p, \\ S\bar{z}^*(\beta) &= \frac{\beta q}{\beta q + (1-\beta)r}(1-(1-\rho)(1-q)), \\ S\underline{z}^*(\beta) &= \frac{\beta(1-q)}{\beta(1-q)+(1-\beta)(1-r)}(1-\rho)q. \end{aligned} \tag{4.6}$$

Here, β is player i's belief that player j is in state w^R at the beginning of period t, the probability that j will choose effort E, and function $a_i z_i^*$ calculates player i's updated belief that player j experiences $E\bar{z}$ at the end of period t while player i experiences $a_i z_i$. Because an action choice of E is from state w^R with certainty, $a_i z_i^*(\beta)$ is player i's starting belief of j in w^R for the following period $t+1$. We calculate the probability of $E\bar{z}$ because players move into rewarding state w^R only after private history $E\bar{z}$. Because the initial state for both players is w^R, $\beta_i^0 = 1$, $i = 1$, 2. The sequence $\{\beta_i^t\}_{t=0}^{\infty}$ is well-defined given the strategy, and each β_i^t is h_i^t-dependent, that is, for different h_i^t, player i may hold different β_i^t.

Associated with each state and given a value of β, there is an expected continuation value corresponding to each action chosen by the players. The value function here calculates *the value* when all players follow the output function $f(\cdot)$. And they are

$$\begin{aligned} V(w^R,\beta) = (1-\delta)\big[\beta u(EE) + (1-\beta)u(ES)\big] \\ + \delta\Big[\big(\beta p + (1-\beta)q\big)V(w^R, E\bar{z}^*(\beta)) + \big(\beta(1-p)+(1-\beta)(1-q)\big)V(w^P, E\underline{z}^*(\beta))\Big] \end{aligned} \tag{4.7}$$

by playing E, and

$$V(w^P,\beta) = (1-\delta)\big[\beta u(SE) + (1-\beta)u(SS)\big] \\ + \delta\Big[\big(\beta q + (1-\beta)r\big)V(w^P, S\bar{z}^*(\beta)) + \big(\beta(1-q) + (1-\beta)(1-r)\big)V(w^P, S\underline{z}^*(\beta))\Big]$$

by playing S.

Assuming that the players follow the output function from the next period on, for whatever value of β, the right hand side of the two equations are the expected payoffs by playing E and S, respectively. For the strategies to constitute an equilibrium, the necessary and sufficient conditions are

$$V(w^R,\beta) \geq V(w^P,\beta) \tag{4.8}$$

for all possible values of β when the player reaches state w^R, and

$$V(w^P,\beta) \geq V(w^R,\beta)$$

for all possible values of β when the player reaches state w^P. This guarantees that the incentives are consistent with the output function $f(\cdot)$.

We rewrite the functional equations (4.7) in a matrix representation. By the consistency between the beliefs over states and those over action choices, we let β denote a player's belief that her opponent will choose effort E. Let V^{EE} denote the continuation promise when both players choose E, and V^{ES}, V^{SE}, and V^{SS} are defined accordingly. We list the one-period outcomes of a player in the order of $E\bar{z}$, $E\underline{z}$, $S\bar{z}$, and $S\underline{z}$. This order is unessential, and a different ordering will cause permutation of row or column vectors of matrices but won't change the fundamental arguments. In matrix form, these equations are written as the following:

$$\begin{aligned}\begin{pmatrix}V^{EE} & V^{ES}\end{pmatrix}\begin{pmatrix}\beta \\ 1-\beta\end{pmatrix} &= (1-\delta)\begin{pmatrix}u(EE) & u(ES)\end{pmatrix}\begin{pmatrix}\beta \\ 1-\beta\end{pmatrix} \\ &+ \delta\Bigg[\begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1\end{pmatrix}\begin{pmatrix}V^{EE} & V^{ES} \\ V^{SE} & V^{SS}\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1\end{pmatrix}\mathbb{P}\begin{pmatrix}1 \\ 0 \\ 0 \\ 0\end{pmatrix} \\ &+ \begin{pmatrix}0 & 1 & 0 & 0\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1\end{pmatrix}\begin{pmatrix}V^{EE} & V^{ES} \\ V^{SE} & V^{SS}\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1\end{pmatrix}\mathbb{P}\begin{pmatrix}0 \\ 1 \\ 0 \\ 0\end{pmatrix}\Bigg],\end{aligned} \tag{4.9}$$

and

$$\begin{aligned}\begin{pmatrix}V^{SE} & V^{SS}\end{pmatrix}\begin{pmatrix}\beta \\ 1-\beta\end{pmatrix} &= (1-\delta)\begin{pmatrix}u(SE) & u(SS)\end{pmatrix}\begin{pmatrix}\beta \\ 1-\beta\end{pmatrix} \\ &+ \delta\Bigg[\begin{pmatrix}0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1\end{pmatrix}\begin{pmatrix}V^{EE} & V^{ES} \\ V^{SE} & V^{SS}\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1\end{pmatrix}\mathbb{P}\begin{pmatrix}0 \\ 0 \\ 1 \\ 0\end{pmatrix} \\ &+ \begin{pmatrix}0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1\end{pmatrix}\begin{pmatrix}V^{EE} & V^{ES} \\ V^{SE} & V^{SS}\end{pmatrix}\begin{pmatrix}1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1\end{pmatrix}\mathbb{P}\begin{pmatrix}0 \\ 0 \\ 0 \\ 1\end{pmatrix}\Bigg].\end{aligned}$$

We explain the major terms in the first equation of (4.9). On the left, we have $V(w^R, \beta) =$

$\beta V^{EE} + (1-\beta)V^{ES}$. The row and column vectors of

$$\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

indicate that player i has experienced $E\bar{z}$ which is the first outcome in the list. The 4×2 and 2×4 matrices are the strategies of player i and j, respectively. For example, player i will assign probability 1 to effort E and 0 to shirking S after $E\bar{z}$, which are given by 1 and 0 in the first row. The matrix $\mathbb{P}$ is a 4×4 probability distribution matrix, in which the first column is i's belief after her effort E over the outcomes of player j together with $E\bar{z}$ of i. Other columns are constructed accordingly. The sum of the entries in the first two columns is 1, so is the sum of those in the last two columns. We have

$$\mathbb{P} = \begin{pmatrix} \beta p\big(1-(1-\rho)(1-p)\big) & \beta(1-p)(1-\rho)p & \cdots \\ \beta p(1-\rho)(1-p) & \beta(1-p)\big(1-(1-\rho)p\big) & \cdots \\ (1-\beta)q\big(1-(1-\rho)(1-q)\big) & (1-\beta)(1-q)(1-\rho)q & \cdots \\ (1-\beta)q(1-\rho)(1-q) & (1-\beta)(1-q)\big(1-(1-\rho)q\big) & \cdots \end{pmatrix}$$

$$\begin{pmatrix} \cdots & \beta q\big(1-(1-\rho)(1-q)\big) & \beta(1-q)(1-\rho)q \\ \cdots & \beta q(1-\rho)(1-q) & \beta(1-q)\big(1-(1-\rho)q\big) \\ \cdots & (1-\beta)r\big(1-(1-\rho)(1-r)\big) & (1-\beta)(1-r)(1-\rho)r \\ \cdots & (1-\beta)r(1-\rho)(1-r) & (1-\beta)(1-r)\big(1-(1-\rho)r\big) \end{pmatrix}.$$

And finally, we view the 2×2 matrix of $V^{a_i a_j}$ as a two-dimensional vector space of linear functions over $\beta \in [0,1]$,

$$V = \begin{pmatrix} \beta V^{EE} + (1-\beta)V^{ES} \\ \beta V^{SE} + (1-\beta)V^{SS} \end{pmatrix}, \tag{4.10}$$

and endow this space with the supreme norm. The right-hand-side of the equations in (4.9) jointly define an operator acting on this vector space, and this operator is solely determined by the strategy profile $\mathfrak{s}$. Let $\mathfrak{s}^*$ denote this operator. The solution of the value functions in (4.7) is the fixed-point of $\mathfrak{s}^*$, and the value functions obtained in this way calculate the expected continuation values if *the players follow the same output functions in perpetuity.* Because $\|\mathfrak{s}^*V - \mathfrak{s}^*V'\| \leq K\|V - V'\|$ for some K, $0 < K < 1$, for general distributions of private signals, $\mathfrak{s}^*$ has a fixed-point, which gives a unique solution of the value functions. This leaves us the consistency of beliefs (4.8) to verify.

We present a detailed solution procedure in Appendix Section C.1, and show possible parameter values such that the grim trigger is sequentially rational. When the discount factor δ is close to 1, the players become sufficiently patient. They may have an incentive to deviate from relentless shirking after a private history of, say, a series of repeated $S\bar{z}$. The grim trigger is no longer sequentially rational in this case.

We state this result in more details. Although we can solve for the value functions $V(w^R, \beta) = \beta V^{EE} + (1-\beta)V^{ES}$ and $V(w^P, \beta) = \beta V^{SE} + (1-\beta)V^{SS}$, the continuation promises V^{ES} and V^{SE} are not credible and cannot be reached. Let β^* denote the point where the two value functions cross. The point divides the range of beliefs $[0, 1]$ into two subintervals $[0, \beta^*)$ and $[\beta^*, 1]$. The beliefs in $[0, \beta^*)$ should sustain the continuation promise V^{SS}, while the beliefs in $[\beta^*, 1]$ should sustain the continuation promise V^{EE}. This consistency is mandated by the condition (4.8).

When the correlation coefficient ρ is close to 0, players receive almost independent private signals. Calculations show that, as functions of belief β,

$$E\bar{z}^*(\beta) > E\underline{z}^*(\beta) > S\bar{z}^*(\beta) > S\underline{z}^*(\beta),$$

which have fixed-points between 0 and 1 also in that order. If the fixed-point of $E\bar{z}^*(\cdot)$ is the only one among them lying in the interval $[\beta^*, 1]$, as the game proceeds, β moves around this fixed-point after $E\bar{z}$, and the grim trigger is sequential.

When both δ and ρ are close to 1, players are patient while private signals are almost perfectly correlated and the monitoring is almost public. In this case, however, $S\bar{z}^*(\beta) > E\underline{z}^*(\beta)$. When some distribution of private signals is chosen such that $\beta^s > \beta^* > \beta^e$, where β^s and β^e are the fixed points of $S\bar{z}^*(\cdot)$ and $E\underline{z}^*(\cdot)$, respectively, after sufficiently many $S\bar{z}$, belief β goes to β^s and will enter the region $[\beta^*, 1]$. It is now optimal to play E rather than perpetual shirking. The grim trigger fails sequential rationality because of the inconsistent beliefs, $S\bar{z}^*\big([0,\beta^*)\big) \nsubseteq [0,\beta^*)$, where the first $[0,\beta^*)$ is the submanifold of i's belief that S by j in punishment state is more likely so i plays S in response which results in $S\bar{z}^*$ and the second $[0,\beta^*)$ is i's belief for the next period such that i is assigned S after $S\bar{z}$ under the grim trigger in response to j's S,[2] but it fails not because of the consecutive outcomes of $S\bar{z}$ which is of little chance of occurrence *a priori*.

4.2 Forgiving Strategies

We now consider a profile that provides incentives to exert effort but is less inclined to punish. Players exert effort after the private signal $\bar{z}$ and shirk after $\underline{z}$, and exert effort in the initial period. In the two-state automaton representation of this profile, the state space, initial state, and output function are the same as the grim trigger. The only difference is the transition function,

$$\tau(w, a_i z_i) = \begin{cases} w^R, & \text{if } z_i = \bar{z}, \\ w^P, & \text{if } z_i = \underline{z}, \end{cases}$$

where a_i, z_i are the player's current period action and signal, respectively. This profile is similar to the forgiving strategies in public monitoring games in that it rewards good signals and punishes bad ones. It requires sufficient correlation between the players' private signals, so that they can coordinate their actions to sustain equilibrium payoffs higher than the stage game Nash equilibrium payoffs.

Let β_i^t, again, denote the probability that player i's belief $\mathfrak{B}_i^t$ assigns to the case of player j being in state w^R when player i observes h_i^t. If β is player i's belief that player j is in state w^R at the beginning of period t, let function $a_i z_i^*$ calculate player i's updated belief that player j experiences $E\bar{z}$ or $S\bar{z}$ at the end of period t while player i experiences $a_i z_i$. We calculate the probability of $E\bar{z}$ and $S\bar{z}$ here because the players move into rewarding state w^R after private history $E\bar{z}$ or $S\bar{z}$. By

[2] There exits some β, $\beta \in [0,\beta^*)$ but $S\bar{z}^*(\beta) \notin [0,\beta^*)$, which is inconsistent because it is now optimal for i to play E rather than S assigned by the grim trigger.

Bayes' rule and (4.5), the functions are given as

$$
\begin{aligned}
E\bar{z}^*(\beta) &= \frac{\beta p}{\beta p + (1-\beta)q}(1-(1-\rho)(1-p)) + \frac{(1-\beta)q}{\beta p + (1-\beta)q}(1-(1-\rho)(1-q)), \\
E\underline{z}^*(\beta) &= \frac{\beta(1-p)}{\beta(1-p)+(1-\beta)(1-q)}(1-\rho)p + \frac{(1-\beta)(1-q)}{\beta(1-p)+(1-\beta)(1-q)}(1-\rho)q, \\
S\bar{z}^*(\beta) &= \frac{\beta q}{\beta q + (1-\beta)r}(1-(1-\rho)(1-q)) + \frac{(1-\beta)r}{\beta q + (1-\beta)r}(1-(1-\rho)(1-r)), \\
S\underline{z}^*(\beta) &= \frac{\beta(1-q)}{\beta(1-q)+(1-\beta)(1-r)}(1-\rho)q + \frac{(1-\beta)(1-r)}{\beta(1-q)+(1-\beta)(1-r)}(1-\rho)r.
\end{aligned} \tag{4.11}
$$

Because the initial state for both players is w^R, $\beta_i^0 = 1$, $i = 1, 2$. The sequence $\{\beta_i^t\}_{t=0}^{\infty}$ is defined by $\beta_i^{t+1} = a_i z_i^*(\beta_i^t)$, and each β_i^t is h_i^t-dependent.

The value functions as functions of belief β in each state are

$$
\begin{aligned}
V(w^R, \beta) &= (1-\delta)\big[\beta u(EE) + (1-\beta)u(ES)\big] \\
&\quad + \delta\Big[\big(\beta p + (1-\beta)q\big)V(w^R, E\bar{z}^*(\beta)) + \big(\beta(1-p) + (1-\beta)(1-q)\big)V(w^P, E\underline{z}^*(\beta))\Big]
\end{aligned} \tag{4.12}
$$

by playing E, and

$$
\begin{aligned}
V(w^P, \beta) &= (1-\delta)\big[\beta u(SE) + (1-\beta)u(SS)\big] \\
&\quad + \delta\Big[\big(\beta q + (1-\beta)r\big)V(w^R, S\bar{z}^*(\beta)) + \big(\beta(1-q) + (1-\beta)(1-r)\big)V(w^P, S\underline{z}^*(\beta))\Big]
\end{aligned}
$$

by playing S. The fact that player i goes into the state of w^R instead of w^P after $S\bar{z}$ is the only difference between the equations here and those in (4.7). As to the matrix representation of (4.12), the majority part of (4.9) remains the same except that player i and j's strategy matrices are changed to

$$
\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}^T \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.
$$

The right-hand-side of the equations in (4.12) are the expected payoffs obtained by playing E and S, respectively, for a given value of β. The value functions calculate the expected continuation values when all players follow the output function $f(\cdot)$. The incentive conditions are

$$V(w^R, \beta) \geq V(w^P, \beta) \tag{4.13}$$

for all possible values of β when the player reaches state w^R, and

$$V(w^P, \beta) \geq V(w^R, \beta)$$

for all possible values of β when the player reaches state w^P.

Before we start to solve the functional equations in (4.12), we digress to examine their connection with imperfect public monitoring games, which are the special case where the correlation coefficient $\rho = 1$ for private monitoring games. When $\rho = 1$, we have $\beta = 1$ if the players are in state w^R, and $\beta = 0$ if the players are in state w^P. The incentive conditions are reduced to

$$V(w^R, \beta = 1) \geq V(w^P, \beta = 1)$$

and

$$V(w^P, \beta = 0) \geq V(w^R, \beta = 0),$$

which are the incentive conditions of the Nash equilibrium for the normal-form game with its payoffs given by the value functions.

Section C.2 in the Appendix shows the solution procedure. Let β^* denote the crossing point of the two value functions, and $\beta^* \in [0, 1]$ separates the range of beliefs $[0, 1]$ into two subintervals. Let β^R denote a lower bound of belief β immediately after outcome $E\bar{z}$ or $S\bar{z}$, and β^P denote an upper bound of belief β immediately after outcome $E\underline{z}$ and $S\underline{z}$. The equilibrium existing condition is $\beta^P \leq \beta^* \leq \beta^R$.

We now examine a numerical example. The calculation is assisted by computer programs, which are all available upon request. Let $\rho = 0.8$, $p = 0.9$, $q = 0.5$, and $r = 0.4$. Given these values and results from (C.8) and (C.9), β^* is an decreasing function with respect to the discount factor δ, and $\delta \in (0.7565, 1)$ guarantees that $\beta^* \in (0,1)$. We turn to the incentive conditions for further restriction on β^*, and δ. The functions $E\bar{z}^*$, $E\underline{z}^*$, $S\bar{z}^*$, and $S\underline{z}^*$ are mild increasing functions of β, and

$$E\bar{z}^*(\beta) > S\bar{z}^*(\beta) > E\underline{z}^*(\beta) > S\underline{z}^*(\beta) \tag{4.14}$$

for any $\beta \in [0,1]$. The condition $\beta^P \leq \beta^*$ is trivially satisfied, because the maximum of $E\underline{z}^*(\cdot)$ and $S\underline{z}^*(\cdot)$ is $E\underline{z}^*(\beta) = 0.18$ when $\beta = 1$ while $\beta^* > 0.18$ for all $\delta \in (0.7565, 1)$, e.g., $\beta^* = 0.4702$ for $\delta = 0.9999$.

For the condition $\beta^* \leq \beta^R$, we search for the lowest possible β after an outcome $S\bar{z}$ because of (4.14). The only way to reach a lower β is the history path $h_i^t = E\underline{z}S\underline{z}\ldots S\underline{z}S\bar{z}$, in which there is a long series of $S\underline{z}$ between $E\underline{z}$ and $S\bar{z}$. We know $\beta_i^0 = 1$ as the initial state is w^R, then $\beta_i^1 = E\underline{z}^*(\beta_i^0) = 0.18$ which is shown above. The fixed point of the function $S\underline{z}^*$ is $\beta^\dagger = 0.0814$. By Lemma 5, $\beta^\dagger = 0.0814$ is the lower bound of belief β after the history consisting of $E\underline{z}$ in period 0 followed by a long series of $S\underline{z}$. Because $S\bar{z}^*(\cdot)$ is an increasing function, we have $\beta^R = S\bar{z}^*(\beta^\dagger) = 0.8820$. When $\delta > 0.7997$, $\beta^* \leq \beta^R$. To summarize, for $\delta \in (0.7997, 1)$, the forgiving strategy is an equilibrium.

A caveat is that, the graphs of function $E\bar{z}^*$, $S\bar{z}^*$ and those of $E\underline{z}^*$, $S\underline{z}^*$ have a difference of length ρ when $\beta = 0$. The strategy profile is sequential only if ρ is close to one, that is, private signals are highly correlated.

4.3 Mixed-strategies

Sequential mixed-strategy profiles behave differently from pure-strategy sequentially rational profiles. For a player to be indifferent about her actions and states, it has to be true that the value functions of different states, as functions of private beliefs, have identical functional form. We use examples close to those studied in Kandori and Obara (2006) as an illustration. The game is the infinitely repeated prisoners' dilemma discussed in the previous sections. We first analyze a strategy profile allowing mixed action choices, but the transition of states is certain.

The strategy profile is symmetric, so the automaton representation is common and we suppress its subscript i. The set of states for player i is $\mathcal{W} = \{w^R, w^P\}$ and the initial state $w_i^0 = w^R$. The output function is given by

$$f(w) = \begin{cases} \alpha \circ E + (1-\alpha) \circ S, & w = w^R, \\ S, & w = w^P, \end{cases}$$

where $\alpha \in (0, 1]$ is the probability of E. The transition function is given by

$$\tau(w, a_i z_i) = \begin{cases} w^R, & \text{if } w = w^R \text{ and either } a_i = E \text{ or } a_i z_i = S\underline{z}, \\ w^P, & \text{otherwise,} \end{cases}$$

where a_i, z_i are the player's current period action and signal, respectively.

We continue to let β_i^t denote the probability that player i's belief $\mathfrak{B}_i^t$ assigns to the case of player j being in *rewarding* state w^R when player i observes h_i^t. By Bayes' rule and (4.5), the function $a_i z_i^*$ calculates player i's updated belief that player j is in state w^R after i experiences $a_i z_i$. We show the detailed derivation for the case when $a_i z_i = E\bar{z}$.

Given β, the probability in player i's ex ante belief that j is in w^R, player i experiences $a_i z_i = E\bar{z}$ and believes that player j will play E with probability $\beta\alpha$ and play S with probability $(1-\beta) + \beta(1-\alpha)$ for the current period. With $a_i z_i = E\bar{z}$, the ex post belief of player i that $a_j z_j = E\bar{z}$, $E\underline{z}$, $S\bar{z}$, and $S\underline{z}$ has the following probabilities,

$$\frac{\beta\alpha p}{\beta\alpha p + \big((1-\beta) + \beta(1-\alpha)\big)q}\big(1 - (1-\rho)(1-p)\big), \tag{4.15}$$

$$\frac{\beta\alpha p}{\beta\alpha p + \big((1-\beta) + \beta(1-\alpha)\big)q}(1-\rho)(1-p),$$

$$\frac{\big((1-\beta) + \beta(1-\alpha)\big)q}{\beta\alpha p + \big((1-\beta) + \beta(1-\alpha)\big)q}\big(1 - (1-\rho)(1-q)\big),$$

and

$$\frac{\big((1-\beta) + \beta(1-\alpha)\big)q}{\beta\alpha p + \big((1-\beta) + \beta(1-\alpha)\big)q}(1-\rho)(1-q).$$

The updated probability that player j is in w^R conditioned on $a_j z_j = E\bar{z}$ and $E\underline{z}$ is 1, while

$$\frac{\beta(1-\alpha)}{(1-\beta)+\beta(1-\alpha)} \circ w^R + \frac{1-\beta}{(1-\beta)+\beta(1-\alpha)} \circ w^P$$

are updated probabilities of player j's states conditioned on $a_j z_j = S\bar{z}$ and $S\underline{z}$. Then, based on the transition function τ, we have

$$E\bar{z}^*(\beta) = \frac{\beta\alpha p + \beta(1-\alpha)q(1-\rho)(1-q)}{\beta\alpha p + \big((1-\beta)+\beta(1-\alpha)\big)q}.$$

In summary, the functions are given as

$$\begin{aligned} E\bar{z}^*(\beta) &= \frac{\beta\alpha p + \beta(1-\alpha)q(1-\rho)(1-q)}{\beta\alpha p + \big((1-\beta)+\beta(1-\alpha)\big)q}, \\ E\underline{z}^*(\beta) &= \frac{\beta\alpha(1-p) + \beta(1-\alpha)(1-q)\big(1-(1-\rho)q\big)}{\beta\alpha(1-p) + \big((1-\beta)+\beta(1-\alpha)\big)(1-q)}, \\ S\bar{z}^*(\beta) &= \frac{\beta\alpha q + \beta(1-\alpha)r(1-\rho)(1-r)}{\beta\alpha q + \big((1-\beta)+\beta(1-\alpha)\big)r}, \\ S\underline{z}^*(\beta) &= \frac{\beta\alpha(1-q) + \beta(1-\alpha)(1-r)\big(1-(1-\rho)r\big)}{\beta\alpha(1-q) + \big((1-\beta)+\beta(1-\alpha)\big)(1-r)}. \end{aligned} \tag{4.16}$$

Because the initial state for both players is w^R, $\beta_i^0 = 1$, $i = 1, 2$. The sequence $\{\beta_i^t\}_{t=0}^{\infty}$ is defined by $\beta_i^{t+1} = a_i z_i^*(\beta_i^t)$, and each β_i^t is h_i^t-dependent as before.

The value functions as functions of belief β in each state are

$$\begin{aligned} V(w^R,\beta) = (1-\delta)\Big[&\alpha\Big(\beta\alpha u(EE) + \big((1-\beta)+\beta(1-\alpha)\big)u(ES)\Big) \\ &+ (1-\alpha)\Big(\beta\alpha u(SE) + \big((1-\beta)+\beta(1-\alpha)\big)u(SS)\Big)\Big] \\ + \delta\Big[&\alpha\Big(\beta\alpha p + \big((1-\beta)+\beta(1-\alpha)\big)q\Big)V\big(w^R, E\bar{z}^*(\beta)\big) \\ &+ \alpha\Big(\beta\alpha(1-p) + \big((1-\beta)+\beta(1-\alpha)\big)(1-q)\Big)V\big(w^R, E\underline{z}^*(\beta)\big) \\ &+ (1-\alpha)\Big(\beta\alpha q + \big((1-\beta)+\beta(1-\alpha)\big)r\Big)V\big(w^P, S\bar{z}^*(\beta)\big) \\ &+ (1-\alpha)\Big(\beta\alpha(1-q) + \big((1-\beta)+\beta(1-\alpha)\big)(1-r)\Big)V\big(w^R, S\underline{z}^*(\beta)\big)\Big] \end{aligned} \tag{4.17}$$

by playing $\alpha \circ E + (1-\alpha) \circ S$, and

$$\begin{aligned} V(w^P, \beta) = {} & (1-\delta)\Big[\beta\alpha u(SE) + \big((1-\beta) + \beta(1-\alpha)\big)u(SS)\Big] \\ & + \delta\bigg[\Big(\beta\alpha q + \big((1-\beta)+\beta(1-\alpha)\big)r\Big)V\big(w^P, S\bar{z}^*(\beta)\big) \\ & \quad + \Big(\beta\alpha(1-q) + \big((1-\beta)+\beta(1-\alpha)\big)(1-r)\Big)V\big(w^P, S\underline{z}^*(\beta)\big)\bigg] \end{aligned} \tag{4.18}$$

by playing S. The equations are highly simplified because of the symmetry in the strategy profiles.

For player i to have incentive in randomizing between E and S, the following two values must be equal: $V(w^R, a_i = E; \beta) = V(w^R, a_i = S; \beta)$.

Simple calculations show that the functional equations have no solutions. This result is due to the over-identification problem. Specifically, there are six effective equations but only five unknowns, so the existence of solutions cannot be guaranteed in general.

In the next example, we introduce one more degree of freedom by allowing randomization in the transition of states. In the infinitely repeated prisoners' dilemma, the set of states for each player, $\mathcal{W} = \{w^R, w^P\}$, the initial states, $w_i^0 = w^R$, and the output functions remain the same,

$$f(w) = \begin{cases} \alpha \circ E + (1-\alpha) \circ S, & w = w^R, \\ S, & w = w^P, \end{cases}$$

where $\alpha \in (0, 1]$ is the probability of E. However, we allow randomization in state transition at w^P, and the transition function is given by

$$\tau(w, a_i z_i) = \begin{cases} w^R, & \text{if } w = w^R \text{ and either } a_i = E \text{ or } a_i z_i = S\bar{z}, \\ w^P, & \text{if } w = w^R \text{ and } a_i z_i = S\underline{z}, \text{ or} \\ & \quad \text{if } w = w^P \text{ and either } a_i = E \text{ or } a_i z_i = S\underline{z}, \\ \theta \circ w^R + (1-\theta) \circ w^P, & \text{if } w = w^P \text{ and } a_i z_i = S\bar{z}, \end{cases}$$

where a_i, z_i are the player's current period action and signal, respectively.

We search for the probabilities, α and θ, so that each player is indifferent between E and S. We continue to let β denote the probability in player i's ex ante belief that j is in w^R. Using the ex post probabilities such as (4.15) and the transition function τ, we derive the belief operators (functions) as

$$\begin{aligned} E\bar{z}^*(\beta) &= \frac{\beta\alpha p + \big((1-\beta)\theta + \beta(1-\alpha)\big)q\big(1-(1-\rho)(1-q)\big)}{\beta\alpha p + \big((1-\beta)+\beta(1-\alpha)\big)q}, \\ E\underline{z}^*(\beta) &= \frac{\beta\alpha(1-p) + \big((1-\beta)\theta + \beta(1-\alpha)\big)(1-q)(1-\rho)q}{\beta\alpha(1-p) + \big((1-\beta)+\beta(1-\alpha)\big)(1-q)}, \\ S\bar{z}^*(\beta) &= \frac{\beta\alpha q + \big((1-\beta)\theta + \beta(1-\alpha)\big)r\big(1-(1-\rho)(1-r)\big)}{\beta\alpha q + \big((1-\beta)+\beta(1-\alpha)\big)r}, \\ S\underline{z}^*(\beta) &= \frac{\beta\alpha(1-q) + \big((1-\beta)\theta + \beta(1-\alpha)\big)(1-r)(1-\rho)r}{\beta\alpha(1-q) + \big((1-\beta)+\beta(1-\alpha)\big)(1-r)}. \end{aligned} \tag{4.19}$$

For player i to have incentive in randomizing between E and S, the following two values must be equal to $V(w^R, \beta)$:

$$\begin{aligned} V(w^R, a_i = E; \beta) = (1-\delta)\Big[\beta\alpha u(EE) + \big((1-\beta)+\beta(1-\alpha)\big)u(ES)\Big] \\ + \delta\Big[\Big(\beta\alpha p + \big((1-\beta)+\beta(1-\alpha)\big)q\Big)V\big(w^R, E\bar{z}^*(\beta)\big) \\ + \Big(\beta\alpha(1-p) + \big((1-\beta)+\beta(1-\alpha)\big)(1-q)\Big)V\big(w^R, E\underline{z}^*(\beta)\big)\Big], \end{aligned} \tag{4.20}$$

and

$$\begin{aligned} V(w^R, a_i = S; \beta) = (1-\delta)\Big[\beta\alpha u(SE) + \big((1-\beta)+\beta(1-\alpha)\big)u(SS)\Big] \\ + \delta\Big[\Big(\beta\alpha q + \big((1-\beta)+\beta(1-\alpha)\big)r\Big)V\big(w^R, S\bar{z}^*(\beta)\big) \\ + \Big(\beta\alpha(1-q) + \big((1-\beta)+\beta(1-\alpha)\big)(1-r)\Big)V\big(w^P, S\underline{z}^*(\beta)\big)\Big]. \end{aligned} \tag{4.21}$$

The equations are little changed because there is no randomization in state transition at w^R.

However, the equation for the value function at state w^P is

$$\begin{aligned} V(w^P,\beta) = {} & (1-\delta)\Big[\beta\alpha u(SE) + \big((1-\beta)+\beta(1-\alpha)\big)u(SS)\Big] \\ & + \delta\Big[\Big(\beta\alpha q + \big((1-\beta)+\beta(1-\alpha)\big)r\Big)\Big(\theta V\big(w^R, S\bar{z}^*(\beta)\big) + (1-\theta)V\big(w^P, S\bar{z}^*(\beta)\big)\Big) \\ & + \Big(\beta\alpha(1-q) + \big((1-\beta)+\beta(1-\alpha)\big)(1-r)\Big)V\big(w^P, S\underline{z}^*(\beta)\big)\Big]. \end{aligned} \tag{4.22}$$

Section C.3 in the Appendix presents the solution. It turns out that in a sequential equilibrium, α is one of the solutions of a quadratic equation. Let us denote it by α^*, then $\theta^* = 1-\alpha^*$. Finally, $V(w^R,\beta) \equiv V(w^P,\beta)$ and they are nontrivial functions of belief β. Because the β coefficient is increasing in α, the player has greater continuation value for a larger α^* given belief β. The equilibrium values depend only on the discount factor, δ, and the effectiveness in distinguishing whether the opponent is shirking, $q-r$, but not on the correlation coefficient ρ.

In solving for probability α of effort E, it can be shown that if $q-r$ is above a threshold, the solutions of the quadratic equation in Section C.3 exist. Moreover, the two players can choose different α^* for their strategy profiles, and their strategies can be asymmetric. A player would have higher continuation values when her opponent choose the larger α^*, that is, her opponent exerting more effort. However, the sum of their payoffs is the largest when both of them choose the larger root of α^*.

4.4 The Set of Strategies

The recovered recursive structure helps us to build functional equations of value functions. The implication of the equations' matrix representation goes beyond mere formality. The set of all possible strategy profiles of a game becomes clear under this representation.

At the beginning of each period, a player forms a belief of her opponents' action choices, either from an initial belief or from Bayesian learning. With her own action choices, she expects continuation values calculated by her value functions, which are entries in a vector of linear functions of beliefs. In the example of the repeated prisoners' dilemma, it's the vector given by (4.10). In a matrix representation, a strategy profile is treated as an operator acting on a vector space of linear value functions endowed with the supreme norm. Regardless in which format a strategy is given, it boils down to the question what probabilities player i will assign to her finite action choices after an outcome, $a_i z_i$, of current period, $i = 1, \ldots, n$. If we put aside the incentive condition of belief consistency, all possible probability-of-action matrices, which may vary from period to period, will exhaust the set of all possible strategy profiles.

To illustrate, we look at the matrix representation of value-function equations in a mixed-strategy profile of the repeated prisoners' dilemma. The representation is simplified because player i only need to solve for the continuation promises V^E and V^S, when she expects her opponent to play E or S. Let β denote the probability of effort E by player j in player i's belief, and we have

$$\begin{pmatrix} V^E & V^S \end{pmatrix} \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} = (1-\delta) \begin{pmatrix} u(EE) & u(ES) \end{pmatrix} \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} + \delta \begin{pmatrix} V^E & V^S \end{pmatrix} \begin{pmatrix} \alpha_j^1 & \alpha_j^2 & \alpha_j^3 & \alpha_j^4 \\ 1-\alpha_j^1 & 1-\alpha_j^2 & 1-\alpha_j^3 & 1-\alpha_j^4 \end{pmatrix} \mathbb{P} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \tag{4.23}$$

and

$$\begin{pmatrix} V^E & V^S \end{pmatrix} \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} = (1-\delta) \begin{pmatrix} u(SE) & u(SS) \end{pmatrix} \begin{pmatrix} \beta \\ 1-\beta \end{pmatrix} + \delta \begin{pmatrix} V^E & V^S \end{pmatrix} \begin{pmatrix} \alpha_j^1 & \alpha_j^2 & \alpha_j^3 & \alpha_j^4 \\ 1-\alpha_j^1 & 1-\alpha_j^2 & 1-\alpha_j^3 & 1-\alpha_j^4 \end{pmatrix} \mathbb{P} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

which are the value functions for player i when she plays E and S, respectively. In this special case of identical value functions, the equations in (4.23) are mathematically equivalent to the form of (4.9) which contains the action matrix of player i. As long as some α_i, the probability of playing E

by i, is strictly between 0 and 1, her value functions for E and S must be identical.

Under the constraint of belief consistency, if we allow probability α of action choices to vary in different periods, the representation will incorporate all types of strategy profiles in the literature, e.g., those with bounded or unbounded recall. As an example, if we go back to the mixed-strategy discussed in the last section, and suppose that the initial state is w^R, the strategy profile is equivalent to always playing $\alpha \circ E + (1 - \alpha) \circ S$ after effort E, always shirking after $S\underline{z}$, and randomizing between E and S after $S\bar{z}$ but with diminishing probabilities in E, i.e., α, $\alpha(1 - \alpha) \ldots$.

The next step is to verify the consistency of beliefs. Given a strategy profile, suppose we have solved its functional equations and obtained the continuation promises, which always exist by the fixed-point argument in the study of the grim trigger strategy. We summarize the verification method in a proposition, which is an equivalent of Proposition 4. Here, the belief operator $a_i z_i^*$ plays a key role in the process.

For simplicity of notations, we look at 2-player games. For $i, j = 1, 2$ and $i \neq j$, let w^{a_j} be a collection of histories for which j is called upon to choose a_j with positive probability. It is allowed to be the whole history space if the strategy is mixed. In each period, player i's belief over w^{a_j} (also a_j), $a_j \in A_j$, is a continuum of probability distributions and forms a manifold denoted by $\mathcal{M}_i$. Let $V^{a_i a_j}$ be the continuation promise when i has played a_i and believes that j has played a_j almost surely. The continuation promise is *credible* if a_i is the best response to the belief that a_j will be played for sure. Some continuation promises may not be credible. In the forgiving strategy, for example, V^{EE} is credible while V^{SE} is not.

Proposition 6. *For a pure-strategy profile to be belief consistent, the following condition is sufficient and necessary. In each period, for $i, j = 1, 2$ and $i \neq j$, player i has a belief manifold $\mathcal{M}_i$ of which there is a partition $\mathcal{P}_i = \{\mathcal{B}_i^{a_j} | a_j \in A_j\}$ where $\mathcal{B}_i^{a_j}$ contains the extremal point of probability 1 over w^{a_j}. If player i is assigned $\hat{a}_i$ after an outcome of $a_i z_i$, then $a_i z_i^*(\mathcal{B}_i^{a_j}) \subseteq \mathcal{B}_i^{\hat{a}_j}$ where $V^{a_i a_j}$ and $V^{\hat{a}_i \hat{a}_j}$ are credible and $\mathcal{B}_i^{\hat{a}_j}$ belongs to a partition of i's beliefs in the next period.*

This condition is trivially satisfied by mixed-strategy profiles because all solved continuation promises are credible and there is no need for the partition of belief manifolds: The submanifold $\mathcal{B}_i^{a_j}$

is $\mathcal{M}_i$ itself for all a_j in the support of the strategy. In summary, players are playing a Bayesian stage game in each period where payoffs are continuation promises and beliefs are updated consistently from last period.

5 Characterizing the Value-function Space

We examine the sets of value functions in sequential equilibria and equilibrium values. A value function of a player is the dot product of her beliefs and continuation promises. The beliefs are discrete probability distributions over her opponents' action choices and the continuation promises are her expected continuation values under those actions for her particular choice of action. Although some continuation promises in pure-strategy profiles are not credible, they are useful in establishing the boundaries in the partition of belief manifolds. Mixed-strategy profiles are special case in which a player's value functions in any period are identical.

In a repeated game with private monitoring where actions and signals are finite, for each player i, $i = 1, \ldots, n$, i has a vector space of value functions, which are linearly defined on her belief manifold. Each $a_i \in A_i$ identifies a value function in each period for i and the vectors are $|A_i|$-dimensional. The vector spaces of all players together form a bundle of spaces, $\mathcal{V}(\delta) \equiv \left\{\mathcal{V}_i(\delta)\right\}_{i=1}^n$, where δ indicates the possible dependence of the discount factor, and we adopt the supreme norm for each space. An element in $\mathcal{V}(\delta)$ is a collection of value-function vectors for player i, $i = 1, \ldots, n$, where an entry of a vector calculates the expected continuation value for certain player by choosing a particular action.

We view an equilibrium profile as actions taken after the outcomes of a period to select continuation promises, which are treated as random variables that assign values to opponents' upcoming action choices with private beliefs of them. Let $\mathcal{W}_i$ be a set collecting functions for player i assigning those continuation promises and $\mathcal{W} \equiv \left\{\mathcal{W}_i\right\}_{i=1}^n$. Abusing notations, let $\mathcal{V}_i$ denote a subset of $\mathbb{R}$ in which a value is calculated by some entry of a vector in $\mathcal{V}_i(\delta)$ as if player i's belief and one of her action choices are given. $\mathcal{V} \equiv \left\{\mathcal{V}_i\right\}_{i=1}^n$. Let $\mathcal{E}(\delta)$ denote the n-dimensional set of equilibrium values, a subset of $\mathcal{V}$. We follow the method in Abreu, Pearce, and Stacchetti (1990) to characterize $\mathcal{E}(\delta)$, and apply *generating function* and the notion of self-generation.

When there are no short-lived players, $\mathbf{A} \equiv \prod_{i=1}^{n} \Delta(A_i)$ is the set of all feasible action profiles. Let $\mathbf{A}_{-i} \equiv \prod_{j \neq i} A_j$ and $\Omega_{-i} \equiv \prod_{j \neq i} \Omega_j$ denote the action space and one-period outcome space of player i's opponents, $1 \leq i, j \leq n$, respectively. Let β_i denote i's private belief over $\mathbf{A}_{-i}$, and it induces a belief over Ω_{-i} given i's action choice and the distribution of private signals. We define random variable $\tilde{\nu}_i^t$ as a function $w_i^t : \mathbf{A}_{-i} \to \mathbb{R}$ with i's period t (prior/initial) private belief β_i^t. Given i and t, there are $|A_i|$ random variables $\tilde{\nu}_i^t$ indexed by a_i, $a_i \in A_i$, and the value functions are in fact $\mathbb{E}[\tilde{\nu}_i^t]$. Define mapping γ_i and Bayesian learning φ_i,

$$\gamma_i : \quad A_i \times Z_i \to \mathcal{W}_i,$$

$$\varphi_i : \quad \mathcal{B}_i^t \times A_i \times Z_i \to \mathcal{B}_i^{t+1},$$

where $\mathcal{W}_i$ is the set of functions w_i, while $\mathcal{B}_i^t$ and $\mathcal{B}_i^{t+1}$ are the sets of player i's beliefs over $\mathbf{A}_{-i}$ at the beginning of period t and $t+1$, respectively. Let γ denote $\{\gamma_i\}_{i=1}^n$ and φ denote $\{\varphi_i\}_{i=1}^n$.

The vector space $\mathcal{V}(\delta)$ for the next period is formed by γ after elements in Ω, the current one-period outcome space for all players. The mapping γ_i maps $a_i z_i$ in period t of player i to a function w_i which assigns values to all a_{-i} in period $t+1$, and the values are player i's expected payoffs from future play ("continuation promises") given her period-$(t+1)$ action choice after $a_i z_i$. Beliefs follow φ. When player i has a prior belief β_i^t over $\mathbf{A}_{-i}$, after experiencing $a_i z_i$, she updates her belief over a_{-i} and then z_{-i}. Expecting her opponents will follow certain strategy and choose an action profile after $a_{-i} z_{-i}$, player i forms belief β_i^{t+1}, the prior belief for period $t+1$.

Definition 13. *Given player i, a mixed action profile $\alpha = (\alpha_i, \alpha_{-i}) \in \mathbf{A}$ is enforceable for player i with initial belief $\beta_i \in \mathcal{B}_i$ if there exists a collection of value assignments $\mathcal{W}_i$, a mapping γ_i, and Bayesian learning φ_i such that for any $a_i' \in A_i$,*

$$\begin{aligned} \mathbf{V}_i(\alpha_i, \beta_i, \gamma_i, \varphi_i) \equiv\ & (1-\delta) u_i(\alpha) + \delta \sum_{a_i} \sum_{z_i} \mathbb{E}^{\alpha, \pi} \Big[\gamma_i(a_i, z_i) \cdot \varphi_i(\beta_i, a_i, z_i) \Big] \\ \geq\ & (1-\delta) u_i(a_i', \alpha_{-i}) + \delta \sum_{z_i} \mathbb{E}^{\alpha_{-i}, \pi} \Big[\gamma_i(a_i', z_i) \cdot \varphi_i(\beta_i, a_i', z_i) \Big], \end{aligned} \tag{5.1}$$

where $\cdot$ is the dot product of vectors and a_i is an action choice in the support of α_i. The pair (γ_i, φ_i) enforces α_i for player i with $\mathcal{W}_i$ and initial belief β_i.

Constructed from belief β_i over $\mathbf{A}_{-i}$, player i believes that the others follow α_{-i}, which may be different from their actual choices because of private monitoring. It is possible that mapping γ_i, $i = 1, \ldots, n$, assigns the same continuation promises to different $a_i z_i$, but the choice of γ_i has ample freedom in an equilibrium.[3] In (5.1) of Definition 13, the dot product is $\mathbb{E}[\tilde{\nu}_i^{t+1}]$ if current period is t. Let ν_i denote $\mathbb{E}[\tilde{\nu}_i]$, which is an element in $\mathcal{V}_i$.

Definition 14. *A payoff $\nu_i \in \mathbb{R}$ is decomposable for player i on $\mathcal{V}_i \subset \mathbb{R}$ if there exists a mixed action profile $\alpha_i \in \Delta(A_i)$ with initial belief $\beta_i \in \mathcal{B}_i$, enforced by (γ_i, φ_i) on $\mathcal{V}_i$, such that*

$$\nu_i \equiv \mathbb{E}[\tilde{\nu}_i] = \mathbf{V}_i(\alpha_i, \beta_i, \gamma_i, \varphi_i).$$

The payoff ν_i is decomposed by the combination $(\alpha_i, \beta_i, \gamma_i, \varphi_i)$ for player i on $\mathcal{V}_i$. Then, $\nu_i = w_i \cdot \beta_i$ for some function w_i. We say that belief β_i supports w_i or ν_i.

For any function w_i, there is a region of supporting beliefs, which yields a range of decomposable payoffs, ν_i. These regions of supporting beliefs form the partition in Proposition 6.

In decomposing payoff vectors $\mathcal{V} \subset \mathbb{R}^n$ for all players, we require consistency of enforceable action profiles and beliefs.

Definition 15. *The payoff vector $\nu \in \mathbb{R}^n$ is consistently decomposed by the combination $(\alpha_i, \beta_i, \gamma_i, \varphi_i)$ of player i on $\mathcal{V}$, $i = 1, 2, \ldots, n$, if each ν_i is decomposed by $(\alpha_i, \beta_i, \gamma_i, \varphi_i)$ and Bayesian learning φ_i is consistent with the mapping γ_i for all i.*

For player i, in moving to the next period, player j has an enforceable action profile for each outcome $a_j z_j$, $j \neq i$. By observing $a_i z_i$, player i updates belief β_i over a_j and then $a_j z_j$. Knowing a menu of j's enforceable action profiles after each $a_j z_j$, player i forms belief $\varphi_i(\beta_i, a_i, z_i)$. Intuitively speaking, updated supporting beliefs should be supporting too. This is exactly the claim of Proposition 6.

[3] There is a case that players choose different action profiles after histories h_i and h_i', $h_i \neq h_i'$ but ending in the same $a_i z_i$. The consistency of beliefs guarantees that this case will not happen.

Definition 16. *Bayesian learning* $\varphi \equiv \{\varphi_i\}_{i=1}^n$ *is consistent with mapping* $\gamma \equiv \{\gamma_i\}_{i=1}^n$ *if, for any* i, $w_i \in \mathcal{W}_i$ *with supporting* β_i, *and any feasible* $a_i \in A_i$, $z_i \in Z_i$, *we have that* $\varphi_i(\beta_i, a_i, z_i)$ *supports* $\gamma_i(a_i, z_i)$.

The function w_i, $i = 1, 2, \ldots, n$, assigning continuation promises to actions, can be written in a normal form representation, which forms a Bayesian game in each period. For example, in two-player games, the value function can be written as $V_i(a_i, \beta_i) = \sum_{a_j \in A_j} V^{a_i a_j} \cdot Prob(a_j)$, where $Prob(a_j)$ is the probability of a_j under β_i. The values $V^{a_i \cdot}$ are the image of w_i. Note that some

Figure 2: The normal form representation of Bayesian games

	$\cdots$	a_j	$\cdots$	a'_j	$\cdots$
$\cdots$					
a_i		$V^{a_i a_j}$, $V^{a_j a_i}$		$V^{a_i a'_j}$, $V^{a'_j a_i}$	
a'_i		$V^{a'_i a_j}$, $V^{a_j a'_i}$		$V^{a'_i a'_j}$, $V^{a'_j a'_i}$	
$\cdots$					

The case of two-player repeated games; for larger number of players, the representation can be extended similarly.

of the continuation promises cannot be reached in a sequential equilibrium because they are not supported by beliefs. The payoffs in the Bayesian games vary from period to period.

We search the set of payoffs that can be decomposed on $\mathcal{V} \subset \mathbb{R}^n$, that is, the set of decomposable $\nu_i = w_i \cdot \beta_i$, $i = 1, \ldots, n$, and rules identifying enforceable action profiles along with enforcing γ. Note that payoffs are continuation promises given by functions $w \in \mathcal{W} \equiv \{\mathcal{W}_i\}_{i=1}^n$, which are associated with enforceable action profiles and sets of supporting beliefs. We approach the characterization of the value-function space and set of equilibrium values through $\mathcal{W}$.

Definition 17. *For* $\mathcal{W} \equiv \left\{\mathcal{W}_i\right\}_{i=1}^n$ *of all players, where* $\mathcal{W}$ *is a collection of functions that map opponents' action choices to real numbers, let* $\Gamma(\mathcal{W}) \equiv \{w_i,\ i = 1, 2, \ldots, n : w_i \cdot \beta_i = \mathbf{V}_i(\alpha_i, \beta_i, \gamma_i, \varphi_i)$ *for some* α_i *enforced by* (γ_i, φ_i) *with* $\mathcal{W}_i$ *and initial belief* β_i*; the combinations* $(\alpha_i, \beta_i, \gamma_i, \varphi_i)$ *are consistent* $\}$, *and let* $\Gamma(\mathcal{V}) \equiv \{\nu_i \in \mathbb{R},\ i = 1, 2, \ldots, n : \nu_i = w_i \cdot \beta_i$ *for all supporting* β_i, *where*

$w_i \in \Gamma(\mathcal{W}_i)$ and $\gamma_i \cdot \varphi_i \in \mathcal{V}_i \subset \mathbb{R}\}$. Define functions for all i

$$\mathbf{Q}_i : \quad \Gamma(\mathcal{W}_i) \to \Delta(A_i),$$

$$\mathbf{B}_i : \quad \Gamma(\mathcal{W}_i) \to \big\{\mathcal{B}_i\big\},$$

and

$$\mathbf{U}_i : \quad \Gamma(\mathcal{W}_i) \to \mathcal{W}_i^{A_i \times Z_i},$$

$$\mathbf{Y}_i : \quad \Gamma(\mathcal{W}_i) \to \big(A_i \times Z_i\big)^*,$$

so that, for all $w_i \in \Gamma(\mathcal{W}_i)$, $i = 1, 2, \ldots, n$, each $\beta_i \in \mathbf{B}_i(w_i)$ supports w_i. $\mathbf{Q}_i(w_i)$ is enforced by $\big(\mathbf{U}_i(w_i), \mathbf{Y}_i(w_i)\big)$ on $\mathcal{W}_i$ with belief β_i, and $\mathbf{V}_i\big(\mathbf{Q}_i(w_i), \beta_i, \mathbf{U}_i(w_i), \mathbf{Y}_i(w_i)\big) = w_i \cdot \beta_i$. The payoff $\nu_i = w_i \cdot \beta_i \in \Gamma(\mathcal{V}_i)$ is decomposed by $\big(\mathbf{Q}_i(w_i), \beta_i, \mathbf{U}_i(w_i), \mathbf{Y}_i(w_i)\big)$ and the decomposition of $\nu \in \Gamma(\mathcal{V})$ is consistent.

Because both $\Gamma(\mathcal{W})$ and $\Gamma(\mathcal{V})$ depend on discount factor δ, we write $\Gamma(\mathcal{W}; \delta)$ and $\Gamma(\mathcal{V}; \delta)$ when necessary. We are interested in $\Gamma(\mathcal{W})$, the collection of continuation promises, which themselves can be supported, that is, $\mathcal{W} \subset \Gamma(\mathcal{W})$. Recall that an element $w \equiv \{w_i\}_{i=1}^n$ in $\mathcal{W} \equiv \big\{\mathcal{W}_i\big\}_{i=1}^n$ is a vector of functions mapping opponents' actions to continuation promises. We endow $\mathcal{W}$ with the topology that is induced by the supreme norm. Because intertemporal averaging is possible in repeated private monitoring game, a convex hull of $\mathcal{W}$, $\text{co}\big(\mathcal{W}\big)$, is a space close under linear combination. We do not allow the use of public correlation device.

Because of the recovered stationary structure of the game, the set of equilibrium payoffs and equilibrium continuation payoffs coincide. The function Γ is the *generating function* for the games. We now introduce the notion of self-generation.

Definition 18. *A set of payoff functions $\mathcal{W} \equiv \big\{\mathcal{W}_i\big\}_{i=1}^n$ is self-generating if $\mathcal{W} \subset \Gamma(\mathcal{W})$.*

The following results[4] on self-generation are a direct extension of their counterparts in public monitoring games, only we need to pay special attention to belief updates and consistency. Players'

[4]All proofs in Section 5 are in the Appendix.

states are marked by their one-period outcomes and we apply the one-shot deviation principle.

Proposition 7. ***Self-generation*** *Suppose there is a universal bound for all* $w \in \mathcal{W}$ *where* $w = \{w_i\}_{i=1}^n$ *is a vector of real-valued functions, if* $\mathcal{W}$ *is self-generating and* $\mathcal{V}$ *is constructed from* $\mathcal{W}$, $\Gamma(\mathcal{V}) \subset \mathcal{E}(\delta)$ *and hence* $\mathcal{V} \subset \mathcal{E}(\delta)$.

Once we have the stationarity structure, certain results related to self-generation and factorization introduced by Abreu, Pearce, and Stacchetti (1990) still hold in private monitoring games. A key difference here is the necessary condition of consistency in beliefs and enforceable action profiles to guarantee the existence of sequential equilibria.

Proposition 8. $\mathcal{E}(\delta) = \Gamma\big(\mathcal{E}(\delta)\big)$.

The essence of the proofs of Proposition 7 and Proposition 8 is the separation of the continuation promises and belief processes, thereby making the equilibrium characterization procedure of private monitoring games close to those under public monitoring. The connection between continuation promises and beliefs is consistency, and the continuation values are the expected continuation promises with given beliefs.

Lemma 3. Γ *is a monotonic operator, that is,* $\mathcal{W} \subset \mathcal{W}'$ *implies* $\Gamma\big(\mathcal{W}\big) \subset \Gamma\big(\mathcal{W}'\big)$.

Lemma 4. *If* $\mathcal{W}$ *is compact,* $\Gamma\big(\mathcal{W}\big)$ *is closed.*

We construct a collection of sets of feasible payoff functions, $\mathcal{W}_i^\dagger$, for $i = 1, \ldots, n$. When player i believes that all players follow action profile $\alpha \in \mathbf{A} \equiv \prod_{i=1}^n \Delta(A_i)$, the set of payoffs i can receive is

$$\mathcal{U}_i \equiv \Big\{u_i(\alpha) \in \mathbb{R} \mid \alpha \in \mathbf{A}\Big\}.$$

The set of *feasible* payoff functions is

$$\mathcal{W}_i^\dagger \equiv \Big\{w_i : \mathbf{A}_{-i} \to \mathcal{U}_i\Big\}.$$

Let $\mathcal{W}^\dagger$ denote $\big\{\mathcal{W}_i^\dagger\big\}_{i=1}^n$. Under the implied topology over the function space with supreme norm, the set $\mathcal{W}^\dagger$ is compact because it is a separable Banach space, the space of functions that map finitely many elements to closed and bounded intervals on $\mathbb{R}$.

We follow the iterative algorithm in Abreu, Pearce, and Stacchetti (1990). The set of feasible payoff functions, $\mathcal{W}^\dagger$, is compact, and payoff functions that can be decomposed over $\mathcal{W}^\dagger$ are also feasible. Hence, $\Gamma\big(\mathcal{W}^\dagger\big) \subset \mathcal{W}^\dagger$. Because $\mathcal{E}(\delta)$ is a fixed point of Γ and Γ is monotonic, if $\mathcal{W}^\dagger$ constructs $\mathcal{V}^\dagger$,

$$\mathcal{E}(\delta) \subset \Gamma^m\big(\mathcal{V}^\dagger\big) \subset \mathcal{V}^\dagger, \quad \forall m.$$

For the decreasing sequence $\big\{\Gamma^m(\mathcal{W}^\dagger)\big\}_m$, define

$$\mathcal{W}^\dagger_\infty \equiv \bigcap_m \Gamma^m\big(\mathcal{W}^\dagger\big).$$

Each $\Gamma^m\big(\mathcal{W}^\dagger\big)$ is compact, and so is $\mathcal{W}^\dagger_\infty$. We have

$$\mathcal{W}^\dagger_\infty \subset \cdots \subset \Gamma^2\big(\mathcal{W}^\dagger\big) \subset \Gamma\big(\mathcal{W}^\dagger\big) \subset \mathcal{W}^\dagger, \tag{5.2}$$

and

$$\mathcal{E}(\delta) \subset \mathcal{V}^\dagger_\infty.$$

We claim that the iterative algorithm computes the set of sequential equilibrium payoffs.

Proposition 9. *$\mathcal{W}^\dagger_\infty$ is self-generating and so is $\mathcal{V}^\dagger_\infty$, and $\mathcal{V}^\dagger_\infty = \mathcal{E}(\delta)$.*

The Proposition implies that the set of sequential equilibrium payoffs $\mathcal{E}(\delta)$ is compact.

Corollary 1. *The set of sequential equilibrium payoffs, $\mathcal{E}(\delta)$, is compact.*

We are not able to obtain monotonicity of payoff set $\mathcal{W}$ with respect to the discount factor δ. It is different from the case of PPE payoffs in public monitoring games mainly because of the requirement of consistent and supporting beliefs. An action profile enforceable under δ_1 may not be supported by existing beliefs for an increased discount factor δ_2, $\delta_1 < \delta_2 < 1$. For example, as players become more patient, strategy profiles involving relentless punishment might not be sequential, so it fails to constitute a sequential equilibrium for very patient players.

6 The Folk Theorem with Private Monitoring

The self-generation method is not practical in characterizing the set of sequential equilibrium payoffs. If you haven't been bored to tears yet, we are ready to discuss the folk theorem for private monitoring games in which the distribution of private signals is general. We are able to characterize the set for 2-player games, and the technique is valid for Nash-threat folk theorem of N-player games, $N \geq 3$, but we are not able to verify the result for those games with player-specific punishment. The folk theorem is partial, which is not surprising as it is unrealistic to expect efficiency when monitoring is private.

If an action profile α is not a stage game equilibrium, for it to be enforceable, deviation must lead to different expected continuation values. We adopt the individual full rank condition, that is, the distribution of player i's signals induced by α should be different from the distribution induced by any profile (α_i', α_{-i}) with $\alpha_i' \neq \alpha_i$.

Definition 19. *Suppose the action space A is finite. The profile α has individual full rank for player i if the $|A_i| \times |Z_i|$ matrix $R_i(\alpha_{-i})$ with elements $\big[R_i(\alpha_{-i})\big]_{a_i z_i} \equiv \pi(z_i|a_i, \alpha_{-i})$ has full row rank. If this holds for all players i, then α has individual full rank.*

If signals are informative, given profile α_{-i} of opponents, player i should expect that certain signal z_i is more likely to occur following a particular action choice a_i. We introduce the concept of *cohort* in characterizing this statistical feature.

Definition 20. *Suppose the action space A is finite. Each action $a_i \in A_i$ of player i has a cohort of private signals at profile $\hat{\alpha} = (\hat{\alpha}_i, \hat{\alpha}_{-i})$ if there exists a partition of Z_i, $\{\tilde{z}_i^{a_i}\}_{a_i \in A_i}$ and $\cup_{a_i \in A_i} \tilde{z}_i^{a_i} = Z_i$, such that the probability of $\tilde{z}_i^{a_i}$ is the maximal entry of the column vector $\{\pi(\tilde{z}_i^{a_i}|a_i', \hat{\alpha}_{-i})\}_{a_i' \in A_i}$ in matrix $R_i(\hat{\alpha}_{-i})$ when $a_i' = a_i$. If this holds true for all player i, then cohorts of private signals exist at profile $\hat{\alpha}$.*

Example. In forming the matrix $R_i(\hat{\alpha}_{-i})$, player i has action choices $A_i = \big\{T, M, B\big\}$ and signals $Z_i = \{\bar{z}, z', \underline{z}\}$. The signals $\bar{z}$, z', and $\underline{z}$ are cohorts of actions T, M, and B, respectively, if $p > q, r$, while $q' > p', r'$ and $1 - r - r' > 1 - p - p', 1 - q - q'$. In this example, $\tilde{z}^T = \{\bar{z}\}$, $\tilde{z}^M = \{z'\}$, and $\tilde{z}^B = \{\underline{z}\}$.

Figure 3: Cohorts of private signals.

	$\bar{z}$	z'	$\underline{z}$
T	p	p'	$1-p-p'$
M	q	q'	$1-q-q'$
B	r	r'	$1-r-r'$

We now examine the set of equilibrium payoffs. We first look at pure-strategy equilibrium profiles and start from their value-function equations in matrix representation:

$$V = (1-\delta)\mathbb{E}\big[u(\alpha)\big] + \delta\mathfrak{s}^*(V),$$

where V is a vector of value functions, $u(\alpha)$ is a matrix of stage game payoffs and the expectation is taken on its row vectors. If action profile $\tilde{a}$ is an inefficient Nash equilibrium, we take partial derivative of both sides of the matrix representation with respect to the belief of $\tilde{a}_i$. What we get is an equation in solving for the coefficient of this belief because the function is linear in beliefs. Because monitoring is imperfect, the coefficient of δ in the second term on the right-hand-side is strictly less than 1, and if players become more patient as discount factor δ goes to 1, the coefficient will converge to 0 and the value functions become flat. If we normalize payoffs on both sides such that $u_i(\tilde{a}) = 0$, it is easy to see that $\|V\|$ goes to zero vector as δ goes to 1. This leads to the following result.

Proposition 10. *(The pure-strategy equilibrium) Suppose there are no short-lived players, A and Z are finite, the range of the feasible payoff functions $\mathcal{W}^\dagger$ in $\mathbb{R}^n$ has nonempty interior, and all the pure-action profiles yielding its extreme points have individual full rank and cohorts of private signals. If $\tilde{v} = (\tilde{v}_1, \ldots, \tilde{v}_n)$ is an inefficient Nash equilibrium and its action profile $\tilde{a}$ has individual full rank and cohorts of private signals. If $\hat{v} = (\hat{v}_1, \ldots, \hat{v}_n)$ of action profile $\hat{a}$ is an interior payoff dominating $\tilde{v}$, and pure-strategy profile built on $\hat{a}$ and $\tilde{a}$ constitute a sequential equilibrium. Then, the equilibrium payoffs converge to $\tilde{v}$ as δ goes to 1.*

We then look at mixed-strategy profiles for efficient equilibrium payoffs. We focus on 2-player games from now on. Similar method can show Nash-threat folk theorem for N-player games, $N \geq 3$, but it requires high-dimensional array operations. We also assume pure-strategy minmaxing for simplicity of exposition.

Proposition 11. *(The mixed-strategy equilibrium) Suppose there are two long-lived players, $i, j = 1, 2$, $i \neq j$, A and Z are finite, the range of the feasible payoff functions $\mathcal{W}^\dagger$ in $\mathbb{R}^2$ has nonempty interior, and all the pure-action profiles yielding its extreme points have individual full rank and cohorts of private signals. If $\tilde{v} = (\tilde{v}_1, \tilde{v}_2)$ is an inefficient minmax payoff and the mutual minmaxing profile $\tilde{a}$ has individual full rank and cohorts of private signals. And $\hat{v} = (\hat{v}_1, \hat{v}_2)$ is an interior payoff dominating $\tilde{v}$. Then, there exist $\underline{\delta} < 1$, $\underline{v} = (\underline{v}_1, \underline{v}_2)$ and $\bar{v} = (\bar{v}_1, \bar{v}_2)$ such that $[\underline{v}_1, \bar{v}_1] \times [\underline{v}_2, \bar{v}_2] \subset \mathcal{E}(\delta)$ for all $\delta \in (\underline{\delta}, 1)$.*

Proof of Proposition 11. Because both action profiles $\hat{a}$ and $\tilde{a}$ have individual full rank and cohorts of private signals, without loss of generality we assume there are two corresponding cohorts of signals $\hat{z}_i$ and $\tilde{z}_i$ for player i, $i = 1, 2$, with the following distributions:

$$\pi(\hat{z}_i | a_i = \hat{a}_i) = \begin{cases} p, & \text{if } a_j = \hat{a}_j, \\ q, & \text{if } a_j = \tilde{a}_j, \end{cases} \tag{6.1}$$

and

$$\pi(\hat{z}_i | a_i = \tilde{a}_i) = \begin{cases} q', & \text{if } a_j = \hat{a}_j, \\ r, & \text{if } a_j = \tilde{a}_j, \end{cases} \tag{6.2}$$

where $0 < q' < p < 1$ and $0 < r < q < 1$.

We order player j's one-period outcomes as $\hat{a}_j\hat{z}_j$, $\hat{a}_j\tilde{z}_j$, $\tilde{a}_j\hat{z}_j$, and $\tilde{a}_j\tilde{z}_j$, and explore all possible mixed-strategy profiles by setting α_j^1, α_j^2, α_j^3, and α_j^4 as probabilities player j would choose $\hat{a}_j$ in the next period after these outcomes. Let $V^{\hat{a}_j}$ and $V^{\tilde{a}_j}$ denote the continuation promises for player i when player j chooses $\hat{a}_j$ and $\tilde{a}_j$ almost surely. And let $(\beta, 1-\beta)$ be the probabilities in player i's belief of j's action choices. The equations in solving for $V^{\hat{a}_j}$ and $V^{\tilde{a}_j}$ in i's value functions can be written as:

$$
\begin{aligned}
\beta V^{\hat{a}_j} + (1-\beta)V^{\tilde{a}_j} &= (1-\delta)\big(\beta u(\hat{a}_i\hat{a}_j) + (1-\beta)u(\hat{a}_i\tilde{a}_j)\big) \\
&+ \delta\Big[\big(\alpha_j^1\beta p + \alpha_j^2\beta(1-p) + \alpha_j^3(1-\beta)q + \alpha_j^4(1-\beta)(1-q)\big)V^{\hat{a}_j} \\
&+ \big((1-\alpha_j^1)\beta p + (1-\alpha_j^2)\beta(1-p) + (1-\alpha_j^3)(1-\beta)q + (1-\alpha_j^4)(1-\beta)(1-q)\big)V^{\tilde{a}_j}\Big],
\end{aligned} \tag{6.3}
$$

and

$$
\begin{aligned}
\beta V^{\hat{a}_j} + (1-\beta)V^{\tilde{a}_j} &= (1-\delta)\big(\beta u(\tilde{a}_i\hat{a}_j) + (1-\beta)u(\tilde{a}_i\tilde{a}_j)\big) \\
&+ \delta\Big[\big(\alpha_j^1\beta q' + \alpha_j^2\beta(1-q') + \alpha_j^3(1-\beta)r + \alpha_j^4(1-\beta)(1-r)\big)V^{\hat{a}_j} \\
&+ \big((1-\alpha_j^1)\beta q' + (1-\alpha_j^2)\beta(1-q') + (1-\alpha_j^3)(1-\beta)r + (1-\alpha_j^4)(1-\beta)(1-r)\big)V^{\tilde{a}_j}\Big].
\end{aligned} \tag{6.4}
$$

In (6.3) setting $\beta = 1$, we have

$$
V^{\hat{a}_j} = (1-\delta)u(\hat{a}_i\hat{a}_j) + \delta\Big[\big(\alpha_j^1 p + \alpha_j^2(1-p)\big)V^{\hat{a}_j} + \big((1-\alpha_j^1)p + (1-\alpha_j^2)(1-p)\big)V^{\tilde{a}_j}\Big],
$$

and rewrite it as

$$
V^{\hat{a}_j} = (1-\delta)u(\hat{a}_i\hat{a}_j) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\big(\alpha_j^1 p + \alpha_j^2(1-p)\big) + \delta V^{\tilde{a}_j}. \tag{6.5}
$$

Similarly, in (6.4) setting $\beta = 1$, and setting $\beta = 0$ in (6.3) and (6.4), respectively, we have

$$
V^{\hat{a}_j} = (1-\delta)u(\tilde{a}_i\hat{a}_j) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\big(\alpha_j^1 q' + \alpha_j^2(1-q')\big) + \delta V^{\tilde{a}_j}, \tag{6.6}
$$

$$
V^{\tilde{a}_j} = (1-\delta)u(\hat{a}_i\tilde{a}_j) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\big(\alpha_j^3 q + \alpha_j^4(1-q)\big) + \delta V^{\tilde{a}_j}, \tag{6.7}
$$

and

$$
V^{\tilde{a}_j} = (1-\delta)u(\tilde{a}_i\tilde{a}_j) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\big(\alpha_j^3 r + \alpha_j^4(1-r)\big) + \delta V^{\tilde{a}_j}. \tag{6.8}
$$

Let (6.5)−(6.6), we have

$$0 = (1-\delta)\big[u(\hat{a}_i\hat{a}_j) - u(\tilde{a}_i\hat{a}_j)\big] + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(p-q')(\alpha_j^1 - \alpha_j^2),$$

and so

$$\alpha_j^1 = \alpha_j^2 - \frac{(1-\delta)\big[u(\hat{a}_i\hat{a}_j) - u(\tilde{a}_i\hat{a}_j)\big]}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(p-q')}. \tag{6.9}$$

Similarly, (6.7)−(6.8),

$$\alpha_j^3 = \alpha_j^4 - \frac{(1-\delta)\big[u(\hat{a}_i\tilde{a}_j) - u(\tilde{a}_i\tilde{a}_j)\big]}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(q-r)}. \tag{6.10}$$

Plugging (6.9) in (6.5), we get

$$V^{\hat{a}_j} = (1-\delta)u(\hat{a}_i\hat{a}_j) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)p\bigg(-\frac{(1-\delta)\big[u(\hat{a}_i\hat{a}_j) - u(\tilde{a}_i\hat{a}_j)\big]}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(p-q')}\bigg) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\alpha_j^2 + \delta V^{\tilde{a}_j},$$

and simplify it to the following

$$V^{\hat{a}_j} = (1-\delta)u(\hat{a}_i\hat{a}_j) - \frac{p}{p-q'}(1-\delta)\big[u(\hat{a}_i\hat{a}_j) - u(\tilde{a}_i\hat{a}_j)\big] + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\alpha_j^2 + \delta V^{\tilde{a}_j}.$$

We solve it for α_j^2,

$$\alpha_j^2 = \frac{\big(V^{\hat{a}_j} - \delta V^{\tilde{a}_j}\big)(p-q') + (1-\delta)\big(q'u(\hat{a}_i\hat{a}_j) - pu(\tilde{a}_i\hat{a}_j)\big)}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(p-q')}.$$

Likewise, plugging (6.10) in (6.7), we get

$$V^{\tilde{a}_j} = (1-\delta)u(\hat{a}_i\tilde{a}_j) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)q\bigg(-\frac{(1-\delta)\big[u(\hat{a}_i\tilde{a}_j) - u(\tilde{a}_i\tilde{a}_j)\big]}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(q-r)}\bigg) + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\alpha_j^4 + \delta V^{\tilde{a}_j},$$

and simplify it to the following

$$V^{\tilde{a}_j} = (1-\delta)u(\hat{a}_i\tilde{a}_j) - \frac{q}{q-r}(1-\delta)\big[u(\hat{a}_i\tilde{a}_j) - u(\tilde{a}_i\tilde{a}_j)\big] + \delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)\alpha_j^4 + \delta V^{\tilde{a}_j}.$$

We solve it for α_j^4,

$$\alpha_j^4 = \frac{(1-\delta)V^{\tilde{a}_j}(q-r) + (1-\delta)\big(ru(\hat{a}_i\tilde{a}_j) - qu(\tilde{a}_i\tilde{a}_j)\big)}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(q-r)}.$$

Finally, using (6.9) and (6.10), we get

$$\begin{aligned}
\alpha_j^1 &= 1 + \frac{1-\delta}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(p-q')}\Big[(p-q')V^{\hat{a}_j} - (1-q')u(\hat{a}_i\hat{a}_j) + (1-p)u(\tilde{a}_i\hat{a}_j)\Big], \\
\alpha_j^2 &= 1 + \frac{1-\delta}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(p-q')}\Big[(p-q')V^{\hat{a}_j} + q'u(\hat{a}_i\hat{a}_j) - pu(\tilde{a}_i\hat{a}_j)\Big], \\
\alpha_j^3 &= \frac{1-\delta}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(q-r)}\Big[(q-r)V^{\tilde{a}_j} - (1-r)u(\hat{a}_i\tilde{a}_j) + (1-q)u(\tilde{a}_i\tilde{a}_j)\Big], \\
\alpha_j^4 &= \frac{1-\delta}{\delta\big(V^{\hat{a}_j} - V^{\tilde{a}_j}\big)(q-r)}\Big[(q-r)V^{\tilde{a}_j} + ru(\hat{a}_i\tilde{a}_j) - qu(\tilde{a}_i\tilde{a}_j)\Big].
\end{aligned} \tag{6.11}$$

For $\alpha_j^{\cdot}$ to be probabilities, a necessary condition is

$$\begin{aligned}
&(p-q')V^{\hat{a}_j} - (1-q')u(\hat{a}_i\hat{a}_j) + (1-p)u(\tilde{a}_i\hat{a}_j) \le 0 \\
&(p-q')V^{\hat{a}_j} + q'u(\hat{a}_i\hat{a}_j) - pu(\tilde{a}_i\hat{a}_j) \le 0 \\
&(q-r)V^{\tilde{a}_j} - (1-r)u(\hat{a}_i\tilde{a}_j) + (1-q)u(\tilde{a}_i\tilde{a}_j) \ge 0 \\
&(q-r)V^{\tilde{a}_j} + ru(\hat{a}_i\tilde{a}_j) - qu(\tilde{a}_i\tilde{a}_j) \ge 0.
\end{aligned} \tag{6.12}$$

Then there exists $\underline{\delta} \in (0,1)$ such that α_j is a well-defined mixed-strategy profile for $\delta \in (\underline{\delta}, 1)$.

If $u(\hat{a}_i\hat{a}_j) \ge u(\tilde{a}_i\hat{a}_j)$, the second inequality is binding, then we have

$$V^{\hat{a}_j} \le u(\hat{a}_i\hat{a}_j) - \frac{p}{p-q'}\Big(u(\hat{a}_i\hat{a}_j) - u(\tilde{a}_i\hat{a}_j)\Big).$$

If $u(\hat{a}_i\hat{a}_j) < u(\tilde{a}_i\hat{a}_j)$, the first inequality is binding, then we have

$$V^{\hat{a}_j} \le u(\hat{a}_i\hat{a}_j) - \frac{1-p}{p-q'}\Big(u(\tilde{a}_i\hat{a}_j) - u(\hat{a}_i\hat{a}_j)\Big).$$

If $u(\hat{a}_i\tilde{a}_j) \geq u(\tilde{a}_i\tilde{a}_j)$, the third inequality is binding, then we have

$$V^{\tilde{a}_j} \geq u(\tilde{a}_i\tilde{a}_j) + \frac{1-r}{q-r}\Big(u(\hat{a}_i\tilde{a}_j) - u(\tilde{a}_i\tilde{a}_j)\Big).$$

If $u(\hat{a}_i\tilde{a}_j) < u(\tilde{a}_i\tilde{a}_j)$, the fourth inequality is binding, then we have

$$V^{\tilde{a}_j} \geq u(\tilde{a}_i\tilde{a}_j) + \frac{r}{q-r}\Big(u(\tilde{a}_i\tilde{a}_j) - u(\hat{a}_i\tilde{a}_j)\Big).$$

Under the condition that $V^{\hat{a}_j} \geq V^{\tilde{a}_j}$, we pick $\underline{v}$ and $\bar{v}$ as the lower bound and upper bound for $V^{\tilde{a}_j}$ and $V^{\hat{a}_j}$, respectively, and this finishes the proof. □

So far we have constructed three types of sequential equilibria in private monitoring games, namely, repeated play of stage game Nash equilibrium, pure-strategy equilibrium, and mixed-strategy equilibrium. Because we treat the games as a sequence of Bayesian stage games with varying payoffs (continuation promises) and evolving consistent beliefs over current action choices of opponents, intertemporal averaging is easy to implement. Let $\text{co}\big(\bigcup \mathcal{W}\big)$ denote the convex hull of the union of the functional vector space pertaining to the equilibria, and let $\text{co}\big(\mathcal{V}\big)$ denote the set of equilibrium values generated by the continuation promises in $\text{co}\big(\bigcup \mathcal{W}\big)$.

Corollary 2. $\mathcal{E}(\delta) = co\big(\mathcal{V}\big)$.

7 Final Remarks

This paper examines sequential equilibria of repeated games with private monitoring, where actions and signals are finite. In characterizing the equilibrium set of a game, we construct a sequence of Bayesian stage games with consecutively updated private beliefs. The payoffs in the Bayesian games are continuation promises which are expected continuation values if players choose an action profile almost surely. To solve for the continuation promises, we build a system of functional equations from a recovered recursive structure. A player's value functions, which are solutions of these equations, are functions of her beliefs over opponents' action choices, and each of her actions identifies a value function. We have fully characterized the equilibrium sets for 2-player games and verified a partial folk theorem, but we need new techniques to investigate N-player games with player-specific punishment.

Appendix

A Glossary of Symbols

$\mathbf{A}$: $\mathbf{A} \equiv \prod_{i=1}^{n} \Delta(A_i)$.

$\mathbf{A}_{-i}$: $\mathbf{A}_{-i} \equiv \prod_{j \neq i} A_j$.

A: $A \equiv \prod_{i=1}^{n} A_i$.

a_i, A_i: Action, and set of actions.

$a_i z_i^*$: A belief operator.

$\mathfrak{B}_i$: A sequence of player i's beliefs over her opponents' states.

$\tilde{\mathfrak{B}}_i$: A sequence of player i's beliefs over her opponents' beliefs.

$\mathcal{B}_i$: $\mathfrak{B}_i^t|_{\mathbf{A}_{-i}}$, for some t.

$\mathbf{B}_i :\ \ \Gamma(\mathcal{W}_i) \to \{\mathcal{B}_i\}$.

$\mathcal{E}(\delta)$: The set of sequential equilibrium payoffs, given discount factor δ.

E, S: Exerting effort, shirking by players.

$\mathcal{F}_i^t$: σ-algebra define on space Ω_i.

$\mathfrak{F}_i$: A filtration of σ-algebras.

f_i: A decision rule of an automaton representation.

$\mathcal{H}$, $\mathcal{H}_i$: A collection of players' private histories, h^t, h_i^t, $t = 1, 2, \ldots, \infty$.

h_i^t: Player i's private history from period 0 to period $t-1$.

$\mathcal{I}(\cdot)$: Statistical inference given the initial belief and histories.

i, j: Identification of players.

$\mathcal{M}$: A manifold, the range of a player's beliefs.

N: The number of intervals in the grid of the numerical analysis of β.

n: The number of long-lived players.

$\mathbb{P}$: A probability matrix in a matrix representation of strategy profiles.

$\mathcal{P}$: A convex and connected partition of $\mathcal{M}$.

p, q, r: The probabilities of good signals given an action profile.

$\mathbf{Q}_i : \quad \Gamma(\mathcal{W}_i) \to \Delta(A_i)$.

$\mathfrak{s}_i$: A behavior strategy of player i.

$\mathfrak{s}^*$: A strategy operator in a matrix representation.

$\mathbf{U}_i : \quad \Gamma(\mathcal{W}_i) \to \mathcal{W}_i^{A_i \times Z_i}$.

$\mathcal{U}_i$: Player i's feasible payoffs.

$u_i^*(z_i, a_i)$: The ex-post payoff of player i after a realization of z_i.

$u_i(a)$: Player i's ex-ante stage game payoff.

$\mathcal{V}_i(\delta)$: The set of value-function vectors.

$\mathcal{V}_i$: The set of continuation values.

$\mathbf{V}_i\big(\alpha_i, \beta_i, \gamma_i, \varphi_i\big)$: The value when $\alpha = (\alpha_i, \alpha_{-i})$ is enforced by γ_i and φ_i.

$V^{a_i a_j}$: A continuation promise.

V: A vector of value functions.

$V_i(w, \beta)$: A function of state w and belief β calculating continuation values.

v_i: A function assigning value to each $\times_{j \neq i} w_j \in \prod_{j \neq i} \mathcal{W}_j$ given w_i.

$\mathcal{W}_i$: The set of states of an automaton representation; or, the set of functions mapping a_{-i} to real numbers.

$\mathcal{W}_i^{\dagger}$: The set of feasible payoff functions.

w_i: A state in $\mathcal{W}_i$; or, a mapping $\mathbf{A}_{-i} \to \mathbb{R}$.

$\mathbf{Y}_i : \quad \Gamma(\mathcal{W}_i) \to \big(A_i \times Z_i\big)^*$.

z_i, Z_i: The private signal observed by player i and the set of signals.

$\Gamma(\cdot)$: The set of functions or payoffs that can be decomposed.

Ω_i: Player i's information space.

α: The probability of effort E in a mixed-strategy profile.

β, β_i, β_i^t: Player i's belief when restricted on opponents' state space, $\mathfrak{B}_i^t|_{\Omega_{-i}}$.

$\gamma_i: \quad A_i \times Z_i \to \mathcal{W}_i$.

η_i: $|\prod_{j \neq i} \mathcal{W}_j| = \eta_i$.

κ_i: $|\mathcal{W}_i| = \kappa_i$.

$\tilde{\nu}_i^t$: A random variable defined as a mapping $\Omega_{-i} \to \mathbb{R}$ with i's period t (prior/initial) private belief.

ν: $\nu = \mathbb{E}[\tilde{\nu}]$.

π: The probability distribution of signals.

ρ: The correlation coefficient of private signals between two players.

$\varphi_i: \quad \mathcal{B}_i^t \times A_i \times Z_i \to \mathcal{B}_i^{t+1}$.

τ_i: A transition function of an automaton representation.

ω_i: An element in Ω_i, the infinite series of player i's actions and signals.

B Illustration: A Two-Period Example

We follow Bhaskar and van Damme (2002) and Mailath and Samuelson (2006) by considering a two-period game, in which the second period represents the implementation of different continuation values with future play in an infinitely repeated game. The same as the one in Mailath and Samuelson (2006), the first-period game is the prisoners' dilemma given on the left in Figure 4, and on the right is the second-period coordination game. There is no discounting. All payoffs are received at the end of the second period, so first-period payoffs are *not* informative of the other player's action choice. We show that when monitoring is private, the choice of E in the first period is possible in an equilibrium under certain conditions.

The monitoring is private. Player i, $i = 1, 2$, observes a *private* signal z_i drawn from the space $\{\underline{z}, \bar{z}\}$. The probability distribution of the private signal vector $(z_1, z_2) = \{\underline{z}, \bar{z}\}^2$ is given by $\pi(z_1 z_2|a)$, where $a \in A$ is the first-period action profile. Let $\bar{z}_i$ denote $z_i = \bar{z}$, and similarly for $\underline{z}_i$. For simplicity, we assume symmetry in the marginal distribution of private signals, and it is given

Figure 4: The two-period example

	E	S
E	2,2	-1,3
S	3,-1	0,0

	G	B
G	3,3	0,0
B	0,0	1,1

The first-period stage game is on the left with the second-period stage game on the right.

by (4.1),

$$\pi(\bar{z}|a) = \begin{cases} p, & \text{if } a = EE, \\ q, & \text{if } a = ES \text{ or } SE, \\ r, & \text{if } a = SS, \end{cases}$$

where $0 < q < p < 1$ and $0 < r < p$. In terms of the joint distribution of private signals, we have combined four equations (4.2),

$$\begin{cases} \pi(\bar{z}_1\bar{z}_2|a) + \pi(\bar{z}_1\underline{z}_2|a) = \pi(\bar{z}_1|a), \\ \pi(\underline{z}_1\bar{z}_2|a) + \pi(\underline{z}_1\underline{z}_2|a) = \pi(\underline{z}_1|a), \\ \pi(\bar{z}_1\bar{z}_2|a) + \pi(\underline{z}_1\bar{z}_2|a) = \pi(\bar{z}_2|a), \\ \pi(\bar{z}_1\underline{z}_2|a) + \pi(\underline{z}_1\underline{z}_2|a) = \pi(\underline{z}_2|a). \end{cases}$$

We continue to use the correlation coefficient ρ,

$$\rho = \frac{\pi(\bar{z}_1\bar{z}_2|a) - \pi^2(\bar{z}|a)}{\pi(\bar{z}|a) - \pi^2(\bar{z}|a)},$$

$\rho \in [0,1]$ where $\rho = 0$, $\rho = 1$ correspond to independent and perfectly correlated, respectively. Solve

(4.2) and the joint distribution of the private signals is given by (4.4)

$$\pi(\bar{z}_1\bar{z}_2|a) = \rho\pi(\bar{z}|a) + (1-\rho)\pi^2(\bar{z}|a),$$
$$\pi(\bar{z}_1\underline{z}_2|a) = \pi(\underline{z}_1\bar{z}_2|a) = (1-\rho)(1-\pi(\bar{z}|a))\pi(\bar{z}|a),$$
$$\pi(\underline{z}_1\underline{z}_2|a) = (1-\pi(\bar{z}|a)) - (1-\rho)(1-\pi(\bar{z}|a))\pi(\bar{z}|a).$$

Player i's *state space* is $\Omega_i \equiv \{E\bar{z}_i, E\underline{z}_i, S\bar{z}_i, S\underline{z}_i\}$, which contains the information of the player's action choice and the corresponding realized signal. Neither player can observe a complete element in $\Omega \equiv \Omega_1 \times \Omega_2$ after the first-period stage game.

We describe the underlying structure for the information update between the first and the second period of the game. Let $\mathcal{F}_i^0 \equiv \{\phi, \Omega_i\}$ denote the trivial σ-algebra defined on Ω_i, where ϕ is the empty set; let $\mathcal{F}_i^1$ be the σ-algebra defined on Ω_i with generating subsets $\{E\bar{z}_i\}$, $\{E\underline{z}_i\}$, $\{S\bar{z}_i\}$, and $\{S\underline{z}_i\}$. Let $\mathcal{F}_1^t \times \mathcal{F}_2^{t'}$ be the product of σ-algebras $\mathcal{F}_1^t$ and $\mathcal{F}_2^{t'}$. For example, the move from $\mathcal{F}_1^0 \times \mathcal{F}_2^0$ to $\mathcal{F}_1^1 \times \mathcal{F}_2^0$ indicates that player 1 has chosen an action for the first-period stage game and observed a private signal, but the players have not yet started the second-period stage game. The sequence $\mathfrak{F}_1 \equiv \{\mathfrak{F}_1^t\}_{t=0}^1$ is a *filtration* of σ-algebras, where $\mathfrak{F}_1^0 = \mathcal{F}_1^0 \times \mathcal{F}_2^0$, $\mathfrak{F}_1^1 = \mathcal{F}_1^1 \times \mathcal{F}_2^0$, and $\mathfrak{F}_1^0 \subseteq \mathfrak{F}_1^1$. The sequence $\mathfrak{F}_2 \equiv \{\mathfrak{F}_2^t\}_{t=0}^1$ is defined similarly. Let $\mathfrak{F}$ denote $\{\mathfrak{F}_i\}_{i=1}^2$.

A behavior strategy for player i is a sequence of functions $\mathfrak{s}_i \equiv \{\mathfrak{s}_i^t\}_{t=0}^1$, which is *adapted* to the filtration $\mathfrak{F}_i$. The adaptiveness is given as follows. The function $\mathfrak{s}_i^0 : \Omega \to \Delta(\{E, S\})$ is measurable with respect to $\mathfrak{F}_i^0$, and also measurable with respect to $\mathfrak{F}_i^1$, where $\Delta(\{E, S\})$ is the set of all mixed-strategies over $\{E, S\}$. The measurability of $\mathfrak{s}_i^0$ means that player i knows what action in $\Delta(\{E, S\})$ to choose at the beginning of the first period, and knows what action has been chosen and the realization of z_i after the first period. The function $\mathfrak{s}_i^1 : \Omega \to \Delta(\{G, B\})$ is measurable with respect to $\mathfrak{F}_i^1$, but *not*, in general, measurable with respect to $\mathfrak{F}_i^0$, which means player i's second-period strategy is a function of her action choice in the first-period and her private signal z_i only. Note $\mathfrak{F}_i^1 = \mathcal{F}_i^1 \times \mathcal{F}_j^0$ (or $\mathcal{F}_j^0 \times \mathcal{F}_i^1$, i, $j = 1, 2$, $i \neq j$), meaning that player i's strategy is private.

For a given strategy profile of players $\mathfrak{s} \equiv \{\mathfrak{s}_i\}_{i=1}^2$, it determines a sequence of players' beliefs $\mathfrak{B} \equiv \{\mathfrak{B}_i\}_{i=1}^2$ with $\mathfrak{B}_i \equiv \{\mathfrak{B}_i^t\}_{t=0}^1$, which are the first-order beliefs in the form of probability distributions over $\Omega = \Omega_1 \times \Omega_2$. Take player i's belief $\mathfrak{B}_i$ for instance, because the strategy profile $\mathfrak{s}$ is given, player i holds an initial (prior) probability distribution $\mathfrak{B}_i^0$ of the elements in Ω before the first-period stage game. For example, $\mathfrak{B}_i^0$ is given by (4.4) for $a = EE$ if $\mathfrak{s}$ specifies that both players choose E in the first period. Immediately before the second-period stage games, $\mathfrak{B}_i^1$ assigns probability one to an element in the subspace Ω_i, while $\mathfrak{B}_i^1|_{\Omega_j}$, the projection of $\mathfrak{B}_i^1$ over the subspace Ω_j, is an updated conditional probability distribution of her opponent's history, e.g., $\pi(a_j z_j = E\bar{z} | a_i z_i = E\bar{z})$, etc.

We define the beliefs of beliefs $\tilde{\mathfrak{B}} \equiv \{\tilde{\mathfrak{B}}_i\}_{i=1}^2$, where $\tilde{\mathfrak{B}}_i \equiv \{\tilde{\mathfrak{B}}_i^t\}_{t=0}^1$, as the second-order beliefs in the form of probability distributions over the opponent's beliefs $\mathfrak{B}_j^t$ for $t = 0$, 1, respectively. $\tilde{\mathfrak{B}}_1^0$, $\tilde{\mathfrak{B}}_2^0$ coincide with $\mathfrak{B}_1^0|_{\Omega_2}$, $\mathfrak{B}_2^0|_{\Omega_1}$ for a given strategy profile $\mathfrak{s}$ because the game is symmetric. After the first period, the construction of the beliefs is straightforward. For example, suppose $\mathfrak{s}$ specifies that both players choose E in the first period, then after the first-period stage game, player i knows Ez_i with certainty, and in $\mathfrak{B}_i^1$, she believes that her opponent is in the state of $a_j z_j$ with probability $\pi(a_j z_j | E z_i)$, which is also the probability of $\mathfrak{B}_j^1(a_j z_j)$, and this probability distribution over $\mathfrak{B}_j^1$ defines $\tilde{\mathfrak{B}}_i^1(E z_i)$, where $a_j \in \{E, S\}$, $z_j \in \{\bar{z}, \underline{z}\}$. Note that $\mathfrak{B}$ are probability distributions over Ω, $\tilde{\mathfrak{B}}_i^1$ is a weighted sum of probability distributions over Ω, and thus itself is also a probability distribution over Ω.

We now describe the strategy profile that supports the choice of E in the first period, and we also provide the necessary and sufficient conditions which impose restrictions on the parameter values, such as p, q, and r. Consider the following strategy profile $\mathfrak{s}$: Player i plays E in the first period, plays G in the second period only when she has played E and observed $\bar{z}$, and plays B in all other cases.

Before we examine the incentive constraints, here are the conditional probabilities for private signals:

$$\pi(\bar{z}_i|a,\bar{z}_j) = 1-(1-\rho)\big(1-\pi(\bar{z}|a)\big),$$
$$\pi(\underline{z}_i|a,\bar{z}_j) = (1-\rho)\big(1-\pi(\bar{z}|a)\big),$$
$$\pi(\bar{z}_i|a,\underline{z}_j) = (1-\rho)\pi(\bar{z}|a),$$
$$\pi(\underline{z}_i|a,\underline{z}_j) = 1-(1-\rho)\pi(\bar{z}|a),$$

where $a = EE$, SE (or ES), SS, and i, $j \in \{1,2\}$, $i \neq j$.

We verify the second-period incentives for two cases, $a = EE$, and $z_i = \bar{z}$, $\underline{z}$. When $a_i z_i = E\bar{z}$, player i receives the payoff of

$$3\pi(\bar{z}_j|a,\bar{z}_i) + 0\pi(\underline{z}_j|a,\bar{z}_i) = 3-3(1-\rho)(1-p)$$

by playing G, and the payoff of

$$0\pi(\bar{z}_j|a,\bar{z}_i) + 1\pi(\underline{z}_j|a,\bar{z}_i) = (1-\rho)(1-p)$$

by playing B. For playing G to be optimal in this case, we need

$$(1-\rho)(1-p) < \frac{3}{4}. \tag{B.1}$$

When $a_i z_i = E\underline{z}$, player i receives the payoff of

$$3\pi(\bar{z}_j|a,\underline{z}_i) + 0\pi(\underline{z}_j|a,\underline{z}_i) = 3(1-\rho)p$$

by playing G, and the payoff of

$$0\pi(\bar{z}_j|a,\underline{z}_i) + 1\pi(\underline{z}_j|a,\underline{z}_i) = 1-(1-\rho)p$$

by playing B. For playing B to be optimal in this case, we need

$$(1-\rho)p < \frac{1}{4}. \tag{B.2}$$

Here, we take the advantage of symmetry and the fact that the probability distribution $\tilde{\mathfrak{B}}_i^1(a_i z_i)$ over $\mathfrak{B}_j^1$ is the same as $\mathfrak{B}_j^1(a_j z_j)|_{\Omega_i}$, which is the probability distribution $\mathfrak{B}_j^1(a_j z_j)$ projected on the space Ω_i, where a_i, $a_j \in \{E, S\}$, $a_i = a_j$, and z_i, $z_j \in \{\bar{z}, \underline{z}\}$, $z_i = z_j$. Note that the second-period incentives only restrict the parameter value of p because for the given strategy $\mathfrak{s}$, Bayesian learning cannot change the belief over $a = EE$ after the first-period stage game.

We now turn to the first-period incentives. Suppose that player j plays E in the first period as specified by the strategy $\mathfrak{s}$. Player i's payoff from effort is

$$\begin{aligned}
&\pi(\bar{z}_i|EE)\ \Big[\ 2 + 3\pi(\bar{z}_j|EE, \bar{z}_i) + 0\pi(\underline{z}_j|EE, \bar{z}_i)\ \Big] \\
&\quad + \pi(\underline{z}_i|EE)\ \Big[\ 2 + 0\pi(\bar{z}_j|EE, \underline{z}_i) + 1\pi(\underline{z}_j|EE, \underline{z}_i)\ \Big] \\
&= p\Big(2 + 3\big[1 - (1-\rho)(1-p)\big]\Big) + (1-p)\Big(2 + \big[1 - (1-\rho)p\big]\Big) \\
&= 5p + 3(1-p) - 4(1-\rho)(1-p)p,
\end{aligned} \tag{B.3}$$

whereas the payoff from shirking is

$$\begin{aligned}
&\pi(\bar{z}_i|SE)\ \Big[\ 3 + 0\pi(\bar{z}_j|SE, \bar{z}_i) + 1\pi(\underline{z}_j|SE, \bar{z}_i)\ \Big] \\
&\quad + \pi(\underline{z}_i|SE)\ \Big[\ 3 + 0\pi(\bar{z}_j|SE, \underline{z}_i) + 1\pi(\underline{z}_j|SE, \underline{z}_i)\ \Big] \\
&= q\Big[3 + (1-\rho)(1-q)\Big] + (1-q)\Big(3 + \big[1 - (1-\rho)q\big]\Big) \\
&= 4 - q.
\end{aligned} \tag{B.4}$$

Note that both players still believe that the other player has played E after the observation of their private signals. For playing E to be optimal in the first period, we need

$$5p + 3(1-p) - 4(1-\rho)(1-p)p > 4 - q. \tag{B.5}$$

The combined restrictions (B.5), (B.1), and (B.2) are the necessary and sufficient conditions for the existence of the equilibrium that supports first-period mutual effort. Because $q < p$ is a strict inequality, (B.5) can be rewritten as

$$\begin{aligned} & 5p + 3(1-p) - 4(1-\rho)(1-p)p > 4 - p \\ \text{or} \quad & -1 + 3p - 4(1-\rho)p + 4(1-\rho)p^2 > 0. \end{aligned} \tag{B.6}$$

From (B.1) and (B.2) we know

$$1 - \frac{3}{4}\frac{1}{1-\rho} < p < \frac{1}{4}\frac{1}{1-\rho}. \tag{B.7}$$

To find the range of ρ that supports the equilibrium, we replace p in the second inequality of (B.6) with the value of $1/4(1-\rho)$. It is easy to show that the equivalent condition is $\rho > 1/2$.

There are two observations of this two-period game. For both players to have incentives to exert effort in the first period, *good* signals have to be highly correlated as is measured by ρ, or equivalently, $\pi(\bar{z}_j|a, \bar{z}_i)$. Although the range of value p is enlarging as ρ increases from 1/2 to 1, value p is *not* crucial for the equilibrium to exist as long as $q < p$, and (B.1) and (B.2) are satisfied.

C Computation and Other Examples

C.1 Grim Trigger

To solve the functional equations in (4.7), we rewrite the value functions in the following form based on Proposition 5,

$$V(w^R, \beta) = \beta V^{EE} + (1-\beta)V^{ES} = A\beta + B$$

and

$$V(w^P, \beta) = \beta V^{SE} + (1-\beta)V^{SS} = C\beta + D,$$

which will simplify some notations. Using (4.6) and the second equation in (4.7), we get

$$\begin{aligned} C\beta + D = &(1-\delta)\big[\beta 3 + (1-\beta)0\big] \\ &+ \delta\Big[\big(\beta q + (1-\beta)r\big)\big(C\frac{\beta q}{\beta q + (1-\beta)r}(1-(1-\rho)(1-q)) + D\big) \\ &+ \big(\beta(1-q) + (1-\beta)(1-r)\big)\big(C\frac{\beta(1-q)}{\beta(1-q)+(1-\beta)(1-r)}(1-\rho)q + D\big)\Big], \end{aligned}$$

which yields

$$C\beta + D = \big(3(1-\delta) + C\delta q\big)\beta + \delta D.$$

This equation holds true for arbitrary $\beta \in [0,1]$ and $\delta < 1$, so by equating the coefficients of the β-terms and the constant terms on both sides, we have

$$V(w^P, \beta) = \frac{3\beta(1-\delta)}{1-\delta q}. \tag{C.1}$$

To solve for $V(w^R, \beta)$, plug (C.1) in the first equation of (4.7):

$$\begin{aligned} A\beta + B = &(1-\delta)\big[\beta 2 + (1-\beta)(-1)\big] \\ &+ \delta\Big[\big(\beta p + (1-\beta)q\big)\big(A\frac{\beta p}{\beta p + (1-\beta)q}(1-(1-\rho)(1-p)) + B\big) \\ &+ \big(\beta(1-p) + (1-\beta)(1-q)\big)\frac{3(1-\delta)}{1-\delta q}\frac{\beta(1-p)}{\beta(1-p)+(1-\beta)(1-q)}(1-\rho)p\Big], \end{aligned}$$

which yields

$$A = \frac{1-\delta}{1-\delta q}\Big(3 + \frac{2\delta(p-q)}{1-\delta p(1-(1-\rho)(1-p))}\Big)$$

and

$$B = -\frac{1-\delta}{1-\delta q}.$$

Equating the two value functions,

$$\frac{1-\delta}{1-\delta q}\Big(3 + \frac{2\delta(p-q)}{1-\delta p(1-(1-\rho)(1-p))}\Big)\beta - \frac{1-\delta}{1-\delta q} = \frac{3\beta(1-\delta)}{1-\delta q},$$

gives the point where the graph of the two functions cross

$$\beta^* = \frac{1 - \delta p(1 - (1-\rho)(1-p))}{2\delta(p-q)}.$$

That is, $V(w^R, \beta) \geq V(w^P, \beta)$ for $\beta \geq \beta^*$ and $V(w^R, \beta) < V(w^P, \beta)$ for $\beta < \beta^*$. So far we haven't restricted the range of β or β^* to be inside $[0, 1]$.

We show the following lemma before investigating the range of β for the two states,

Lemma 5. *Let $\beta^\dagger$ be the fixed point of the function $a_i z_i^*$. If $a_i z_i^*$ is an increasing function, and $\beta_i^s \geq \beta^\dagger$, then the sequence $\{\beta_i^t\}_{t=s}^\infty$, given by $\beta_i^{t+1} = a_i z_i^*(\beta_i^t)$, has a lower bound $\beta^\dagger$. Likewise, if $\beta_i^s \leq \beta^\dagger$, $\beta^\dagger$ is the upper bound of the sequence.*

Proof. We only show the case when $\beta_i^s \geq \beta^\dagger$.

Shown by induction. The claim holds true for s, and we assume $\beta_i^t \geq \beta^\dagger$ for some t. Let $a_i z_i^*$ acting on both sides of the inequality. Because the function is monotonic and $\beta^\dagger$ is its fixed point, we have

$$a_i z_i^*(\beta_i^t) \geq a_i z_i^*(\beta^\dagger)$$

and

$$\beta_i^{t+1} \geq \beta^\dagger.$$

□

At the beginning of period t, the only history path that leads to state w^R is a series of $E\bar{z}$ from period 0 to period $t-1$, $h_i^t = E\bar{z}E\bar{z}\ldots E\bar{z}$. $E\bar{z}^*(\cdot)$ is an increasing function and $\beta_i^0 = 1$, which implies that the sequence $\{\beta_i^t\}_{t=0}^\infty$ given by $\beta_i^{t+1} = E\bar{z}^*(\beta_i^t)$ has the fixed point of $E\bar{z}^*(\cdot)$ as its lower bound. To solve for the fixed point β^R of $E\bar{z}^*(\cdot)$, we let

$$\beta = \frac{\beta p}{\beta p + (1-\beta)q}(1 - (1-\rho)(1-p)),$$

which yields

$$\beta^R = \max\left\{0, \quad \frac{p(1-(1-\rho)(1-p)) - q}{p-q}\right\}.$$

For state w^P at the beginning of period t, we look for the upper bound of the belief β. The action and signal combination that can lead to higher β other than $E\bar{z}$ is $S\bar{z}$. So the history in which $E\underline{z}$ in period 0 and $S\bar{z}$ afterwards leads to the highest possible β. The fixed point of $S\bar{z}^*(\cdot)$ is

$$\beta = 0 \text{ and } \frac{q(1-(1-\rho)(1-q))-r}{q-r},$$

and the belief immediately after period 0 $E\underline{z}$ is $(1-\rho)p$. If $q < r$ or $q > r$ but very close, the second fixed point is outside $[0,1]$, $\beta^P = (1-\rho)p$. Otherwise, if the point is inside $[0,1]$,

$$\beta^P = \max\left\{(1-\rho)p, \quad \frac{q(1-(1-\rho)(1-q))-r}{q-r}\right\}.$$

Finally, the equilibrium conditions are summarized as $\beta^P \leq \beta^* \leq \beta^R$. Let us look at a numerical example, where $\rho = 0.75$, $p = 0.8$, $q = 0.6$, and $r = 0.4$. For $\delta \in [0.9259, 0.9615]$, the grim trigger is an equilibrium. When the discount factor, δ, is getting close to 1, a player becomes sufficiently patient. She may have an incentive to deviate from relentless shirking after a private history of, say, a series of repeated $S\bar{z}$. The grim trigger is no longer sequentially rational.

There are other cases that the grim trigger fails to constitute an equilibrium. For example, when $p - q = \varepsilon > 0$, and ε is very small, we have $\beta^* > 1$ and $\beta^R = 0$.

C.2 Forgiving Strategies

To solve the functional equations in (4.12), we continue to use the following form,

$$V(w^R, \beta) = \beta V^{EE} + (1-\beta)V^{ES} = A\beta + B$$

and

$$V(w^P, \beta) = \beta V^{SE} + (1-\beta)V^{SS} = C\beta + D.$$

Using (4.11), we get

$$
\begin{aligned}
A\beta + B = {} & (1-\delta)\big[\beta 2 + (1-\beta)(-1)\big] \\
& + \delta\Big[A\beta p\big(1-(1-\rho)(1-p)\big) + A(1-\beta)q\big(1-(1-\rho)(1-q)\big) \\
& + B\big(\beta p + (1-\beta)q\big) + C\beta(1-p)(1-\rho)p \\
& + C(1-\beta)(1-q)(1-\rho)q + D\big(\beta(1-p) + (1-\beta)(1-q)\big)\Big]
\end{aligned}
\tag{C.2}
$$

and

$$
\begin{aligned}
C\beta + D = {} & (1-\delta)\big[\beta 3 + (1-\beta)0\big] \\
& + \delta\Big[A\beta q\big(1-(1-\rho)(1-q)\big) + A(1-\beta)r\big(1-(1-\rho)(1-r)\big) \\
& + B\big(\beta q + (1-\beta)r\big) + C\beta(1-q)(1-\rho)q \\
& + C(1-\beta)(1-r)(1-\rho)r + D\big(\beta(1-q) + (1-\beta)(1-r)\big)\Big].
\end{aligned}
\tag{C.3}
$$

Directly solving (C.2) and (C.3) is tedious but possible. Fortunately, we are only interested in the signs of $B-D$ and $(A-C)+(B-D)$, and they are $V^{ES}-V^{SS}$ and $V^{EE}-V^{SE}$. These values are (C.2)−(C.3) when $\beta = 0$, 1, and a necessary condition requires their signs to be negative and positive, respectively. Let (C.2)−(C.3),

$$
\begin{aligned}
(A-C)\beta + (B-D) = {} & -(1-\delta) + \delta\Big[(A+B-D)\Big(\big(\beta p + (1-\beta)q\big) - \big(\beta q + (1-\beta)r\big)\Big) \\
& - (A-C)\Big(\beta(1-\rho)p(1-p) + (1-\beta)(1-\rho)q(1-q) \\
& - \beta(1-\rho)q(1-q) - (1-\beta)(1-\rho)r(1-r)\Big)\Big].
\end{aligned}
\tag{C.4}
$$

To get rid of A, we equate coefficients of β-terms on both sides of (C.2),

$$
\begin{aligned}
A = {} & 3(1-\delta) + \delta\Big[(A+B-D)(p-q) \\
& - (A-C)(1-\rho)\big(p(1-p) - q(1-q)\big)\Big].
\end{aligned}
\tag{C.5}
$$

In (C.4), set $\beta = 0$,

$$B - D = -(1-\delta) + \delta\Big[(A + B - D)(q - r) - (A - C)(1-\rho)\big(q(1-q) - r(1-r)\big)\Big], \tag{C.6}$$

and set $\beta = 1$,

$$(A - C) + (B - D) = -(1-\delta) + \delta\Big[(A + B - D)(p - q) - (A - C)(1-\rho)\big(p(1-p) - q(1-q)\big)\Big]. \tag{C.7}$$

So $A = (A - C) + (B - D) + 4(1-\delta)$. Let $X = (A - C) + (B - D)$ and $Y = B - D$, then $A = X + 4(1-\delta)$. From (C.6) and (C.7),

$$Y = -(1-\delta) + \delta\Big[\big(X + Y + 4(1-\delta)\big)(q - r) - (X - Y)(1-\rho)\big(q(1-q) - r(1-r)\big)\Big],$$

and

$$X = -(1-\delta) + \delta\Big[\big(X + Y + 4(1-\delta)\big)(p - q) - (X - Y)(1-\rho)\big(p(1-p) - q(1-q)\big)\Big].$$

Rewriting the above two equations, we get

$$\delta\Big((q-r) - (1-\rho)\big(q(1-q) - r(1-r)\big)\Big)X - \Big(1 - \delta(q-r) - \delta(1-\rho)\big(q(1-q) - r(1-r)\big)\Big)Y = (1-\delta) - 4(q-r)\delta(1-\delta), \tag{C.8}$$

and

$$\Big(-1 + \delta(p-q) - \delta(1-\rho)\big(p(1-p) - q(1-q)\big)\Big)X + \delta\Big((p-q) + (1-\rho)\big(p(1-p) - q(1-q)\big)\Big)Y$$
$$= (1-\delta) - 4(p-q)\delta(1-\delta). \tag{C.9}$$

For the graphes of the two value functions to cross at $\beta^* \in [0,1]$, the necessary and sufficient conditions are $X \geq 0$ and $Y \leq 0$ with at least one strict inequality. Then $\beta^* = -Y/(X-Y)$. Let β^R denote a lower bound of belief β immediately after outcome $E\bar{z}$ or $S\bar{z}$, and β^P denote a upper bound of belief β immediately after outcome $E\underline{z}$ and $S\underline{z}$. Finally, the equilibrium existing condition is $\beta^P \leq \beta^* \leq \beta^R$.

C.3 Mixed-strategies

We continue to assume that the value functions are in the following form,

$$V(w^R, \beta) = A\beta + B$$

and

$$V(w^P, \beta) = C\beta + D.$$

Solving equation (4.20), we have

$$A = \frac{3\alpha(1-\delta)}{1-\delta\big(\alpha + (1-\alpha-\theta)q\big)} \quad \text{and} \quad B = -1 + \frac{3\alpha\theta q\delta}{1-\delta\big(\alpha + (1-\alpha-\theta)q\big)}.$$

Expanding (4.20)$-$(4.21) with (4.19), we have

$$\begin{aligned} 0 = &-(1-\delta) + \delta\Big[\big((1-\beta)\theta + \beta(1-\alpha)\big)(q-r)A \\ &+ \Big(\beta\alpha(1-q) + \big((1-\beta)\theta + \beta(1-\alpha)\big)(1-r)(1-\rho)r\Big)\big(A-C\big) \\ &+ \Big(\beta\alpha(1-q) + \big((1-\beta) + \beta(1-\alpha)\big)(1-r)\Big)\big(B-D\big)\Big]. \end{aligned}$$

Then,

$$\Big(\alpha(1-q)+(1-\alpha-\theta)r(1-\rho)(1-r)\Big)\big(A-C\big)-\alpha(q-r)\big(B-D\big)$$
$$=-(1-\alpha-\theta)(q-r)A, \tag{C.10}$$
$$\delta\theta r(1-\rho)(1-r)\big(A-C\big)+\delta(1-r)\big(B-D\big)=(1-\delta)-\delta\theta(q-r)A. \tag{C.11}$$

Likewise, let (4.21)−(4.22), we have

$$\begin{aligned}\big(A-C\big)\beta+\big(B-D\big)=&\delta(1-\theta)\Big(\beta\alpha q+\big((1-\beta)\theta+\beta(1-\alpha)\big)r\big(1-(1-\rho)(1-r)\big)\Big)\big(A-C\big)\\&+\delta(1-\theta)\Big(\beta\alpha q+\big((1-\beta)+\beta(1-\alpha)\big)r\Big)\big(B-D\big).\end{aligned}$$

Then,

$$\Big[1-\delta(1-\theta)\Big(\alpha q+(1-\alpha-\theta)r\big(1-(1-\rho)(1-r)\big)\Big)\Big]\big(A-C\big)=\delta\alpha(1-\theta)(q-r)\big(B-D\big), \tag{C.12}$$

$$\delta(1-\theta)\theta r\big(1-(1-\rho)(1-r)\big)\big(A-C\big)=\big(1-\delta(1-\theta)r\big)\big(B-D\big). \tag{C.13}$$

We treat A, B as functions of α and θ, and then solve equations (C.10), (C.11), (C.12), and (C.13) to identify $A-C$, $B-D$, α, and θ.

Solving equation (C.12)−$\delta(1-\theta)$(C.10), we have

$$A-C=\frac{\delta(1-\theta)(1-\alpha-\theta)(q-r)A}{1-\delta(1-\theta)\big(\alpha+(1-\alpha-\theta)r\big)}.$$

Solving equation (C.13)+$(1-\theta)$(C.11), we have

$$B-D=-\frac{(1-\delta)(1-\theta)}{1-\delta(1-\theta)}+\frac{\delta(1-\theta)\theta r}{1-\delta(1-\theta)}\big(A-C\big)+\frac{\delta(1-\theta)\theta(q-r)}{1-\delta(1-\theta)}A.$$

Further calculation yields little insight but messy expressions. Note that either {(C.10), (C.11)} or {(C.12), (C.13)} is still independent constraint of α and θ. The numerical solutions of many examples implies that $\alpha+\theta = 1$, so we conjecture that this constraint holds true in general solutions. Thus,

$$A = \frac{3\alpha(1-\delta)}{1-\delta\alpha}, \quad B = -1 + \frac{3\alpha(1-\alpha)q\delta}{1-\delta\alpha},$$

and

$$A - C = 0, \quad B - D = -\frac{(1-\delta)\alpha}{1-\delta\alpha} + \frac{\delta\alpha(1-\alpha)(q-r)}{1-\delta\alpha}A.$$

If we impose $B - D = 0$, then {(C.12), (C.13)} is automatically satisfied. From $B - D = 0$, we have

$$3\delta(q-r)\alpha^2 - 3\delta(q-r)\alpha - \delta\alpha + 1 = 0.$$

Let α^* denote one of the solutions of this quadratic equation if at least one of the solutions exists and lies in the interval $(0, 1]$. For the solutions to exist, we need

$$\Delta \equiv \delta\big(9\delta(q-r)^2 + 6(\delta-2)(q-r) + \delta\big) \geq 0.$$

Let $b^- = (2 - \delta - 2\sqrt{1-\delta})/3\delta$ and $b = (2 - \delta + 2\sqrt{1-\delta})/3\delta$ when δ is sufficiently close to 1. For the Δ to be nonnegative, we need $q - r \leq b^-$ or $q - r \geq b$. For the solutions of α to be in $(0, 1]$, a necessary condition is that the lowest extreme of the quadratic curve for α defined by the quadratic equation is between 0 and 1,

$$0 \leq \frac{3\delta(q-r) + \delta}{6\delta(q-r)} \leq 1,$$

which implies $q - r \geq 1/3$. Because $b \to 1/3+$ as $\delta \to 1$, $q - r \geq b$ is a necessary restriction on the probability distributions of the private signals for the solutions of α to exist. When such α^* exists, let $\theta^* = 1 - \alpha^*$. Finally, $V(w^R, \beta) \equiv V(w^P, \beta)$, and they are nontrivial as functions of the private belief β. Because A is increasing in α, given a private belief β, the player has greater continuation value for larger α^*. In any case, the equilibrium values depend on the discount factor, δ, and the likelihood to distinguish whether the opponent is shirking, $q - r$, but they are unrelated to the correlation parameter ρ.

C.4 Three-player Coordination Game

In this section, we show a nontrivial pure-strategy sequential equilibrium in a three-player coordination game. This example illustrates two facts: first, coordination is possible in private monitoring games even when the private signals are conditionally independent; and second, the specification of belief-dependent action choices is unnecessary.

Borrowed from Fudenberg and Maskin (1986), the example is a repeated play of the three-player coordination game given in Figure 5. Players i, j, and $k \in \{1, 2, 3\}$ value group activities. They each receive payoff 1 if all of them can coordinate their movements by playing *Forward* (F) or playing *Backward* (B) together, and they receive payoff 0 otherwise. In the version of repeated games, the payoffs shown in the matrices are ex ante.

Figure 5: The three-player game

	F	B
F	1,1,1	0,0,0
B	0,0,0	0,0,0

	F	B
F	0,0,0	0,0,0
B	0,0,0	1,1,1

The third player chooses the left matrix by playing F and the right matrix by playing B.

The monitoring technology is privately observed signals which are conditional independent, and for simplicity, we assume the signal distributions are symmetric. Player i, $i \in \{1, 2, 3\}$, observes a *private* signal z_i drawn from the space $\{\underline{z}, \bar{z}\}$. The probability distribution of the private signal vector $(z_1, z_2, z_3) = \{\underline{z}, \bar{z}\}^3$ is given by $\pi(z_1 z_2 z_3 | a)$, where $a \in A$ is the stage game action profile. Let $\bar{z}_i$ again denote $z_i = \bar{z}$, and similarly for $\underline{z}_i$. The marginal distribution of i's private signals is given by (C.14),

$$\pi(\bar{z}_i | a) = \begin{cases} p, & \text{if } a_j a_k = FF, \\ q, & \text{if } a_j a_k = FB \text{ or } BF, \\ r, & \text{if } a_j a_k = BB, \end{cases} \tag{C.14}$$

where $0 < q < p < 1$ and $0 < r < p$.

There are two equilibrium strategy profiles, in which all players repeatedly play stage game Nash equilibria, (F,F,F) or (B,B,B). The equilibria are efficient in offering the highest payoffs possible for the players. We now examine a nontrivial pure-strategy equilibrium profile.

In each period, a player, i, considers possible action choices by other players j and k, which are represented by i's belief β_i over outcomes (F,F), (B,B), (F,B) or (B,F) of j, k. Let $\beta_i = (\beta^{(0)}, \beta^{(1)}, \beta^{(2)})$ be the probability distribution over outcomes in which how many B are played in current period. Note that (F,B) and (B,F) cause no difference because of symmetry. At first glance, a pure-strategy profile may specify belief-dependent action choices, which could be a strategy of countably infinite states at most because finite-period history paths are countable. We will show that this is not necessary.

Suppose that the pure-strategy is belief-dependent, and we now derive its equivalent automaton representation. Any player i will divide all of her beliefs into two disjoint subsets, $\mathcal{W} = \{w^F, w^B\}$, which specify her action choices F and B, respectively. For continuation of play, the classification of the beliefs of i's opponents, j and k, matters. Abusing notations, let $\beta_i = (\beta^{(0)}, \beta^{(1)}, \beta^{(2)})$ denote i's beliefs over her opponents' beliefs. Given i's initial belief, and for any of her finite-period history, belief operators $a_i z_i^*$, $a_i \in \{F, B\}$, $z_i \in \{\bar{z}, \underaccent{\bar}{z}\}$, uniquely identify interim beliefs, which belong to either w^F or w^B. By Proposition 1, $\mathcal{W}_i$, β_i are sufficient statistics for continuation of play.

We look at the belief operators first. The belief β_i^t is a vector $(\beta^{(0)}, \beta^{(1)}, \beta^{(2)})$ denoting, in player i's belief, the probabilities assigned to players j and k's classification of their beliefs. We use the operator $F\bar{z}^*$ as an example to illustrate the procedure in which player i revises her belief from β_i^t to β_i^{t+1} after period t outcome $F\bar{z}$. The β_i^t is player i's initial belief of the combined action choices of players j and k. After experiencing $F\bar{z}$, player i updates her belief of action choices (F,F), $(F,B)/(B,F)$, and (B,B) with probabilities

$$\frac{\beta^{t,(0)}p}{\beta^{t,(0)}p + \beta^{t,(1)}q + \beta^{t,(2)}r}, \quad \frac{\beta^{t,(1)}q}{\beta^{t,(0)}p + \beta^{t,(1)}q + \beta^{t,(2)}r}, \quad \frac{\beta^{t,(2)}r}{\beta^{t,(0)}p + \beta^{t,(1)}q + \beta^{t,(2)}r},$$

respectively. The probability of signal realization $(\bar{z}, \bar{z})$ for j, k in player i's belief is

$$\frac{\beta^{t,(0)}p \cdot p^2 + \beta^{t,(1)}q \cdot pq + \beta^{t,(2)}r \cdot q^2}{\beta^{t,(0)}p + \beta^{t,(1)}q + \beta^{t,(2)}r},$$

and the other probabilities are calculated similarly.

Because $\mathcal{W} = \{w^F, w^B\}$ is a partition of player i's belief manifold, $a_i z_i^*(\beta_i) \in w^F$, or w^B, for $a_i \in \{F, B\}$, $z_i \in \{\bar{z}, \underline{z}\}$. By construction of $\mathcal{W}$, the belief-dependent strategy profile induce a transition function

$$\tau(w, a_i z_i) = \begin{cases} w^F, & \text{if } a_i z_i^*(\beta) \in w^F \text{ for } \beta \in w, \\ w^B, & \text{if } a_i z_i^*(\beta) \in w^B \text{ for } \beta \in w. \end{cases}$$

The transition function is well defined because we cannot have the case in which, for β, $\beta' \in w$, and some a_i, z_i, $a_i z_i^*(\beta) \in w^F$ but $a_i z_i^*(\beta') \in w^B$, as geometrically, $a_i z_i^*$ is a rotation operation acting on belief vectors such that $a_i z_i^*(w) \subseteq w$, $w = w^F$, or w^B. We thus have an automaton representation of a pure-strategy profile.

We now examine a forgiving-type strategy profile, which is coordinating by private signals. The strategy has a simple two-state automaton representation. The state space is $\mathcal{W} = \{w^F, w^B\}$, where w^F denotes the state of moving forward and w^B denotes the state of moving backward. The output function is

$$f(w^F) = F, \quad \text{and} \quad f(w^B) = B,$$

and the transition function is

$$\tau(w, a_i z_i) = \begin{cases} w^F, & \text{if } z_i = \bar{z}, \\ w^B, & \text{if } z_i = \underline{z}, \end{cases}$$

where a_i, z_i are the player's current period action and signal, respectively. Player i's initial state is $\mu_i = \mu_i^F \circ w^F + \mu_i^B \circ w^B$, a probability distribution over $\{w^F, w^B\}$, and initial belief is β_i^0, which is consistent with μ_j and μ_k.

The value functions as functions of belief β in each state are

$$\begin{aligned} V(w^F, \beta) = (1-\delta)\big[\beta^{(0)} u(FFF) + \beta^{(1)} u(FFB) + \beta^{(2)} u(FBB)\big] \\ + \delta\Big[\big(\beta^{(0)} p + \beta^{(1)} q + \beta^{(2)} r\big) V(w^F, F\bar{z}^*(\beta)) \\ + \big(\beta^{(0)}(1-p) + \beta^{(1)}(1-q) + \beta^{(2)}(1-r)\big) V(w^B, F\underline{z}^*(\beta))\Big] \end{aligned} \tag{C.15}$$

by playing F, and

$$\begin{aligned} V(w^B,\beta) &= (1-\delta)\big[\beta^{(0)}u(BFF)+\beta^{(1)}u(BFB)+\beta^{(2)}u(BBB)\big] \\ &\quad + \delta\Big[\big(\beta^{(0)}p+\beta^{(1)}q+\beta^{(2)}r\big)V(w^F,B\bar{z}^*(\beta)) \\ &\quad + \big(\beta^{(0)}(1-p)+\beta^{(1)}(1-q)+\beta^{(2)}(1-r)\big)V(w^B,B\underline{z}^*(\beta))\Big] \end{aligned}$$

by playing B. Again, we conjecture that the value functions are in the following form,

$$V(w^F,\beta) = f^{(0)}\beta^{(0)} + f^{(1)}\beta^{(1)} + f^{(2)}\beta^{(2)},$$

and

$$V(w^B,\beta) = b^{(0)}\beta^{(0)} + b^{(1)}\beta^{(1)} + b^{(2)}\beta^{(2)}.$$

The functional equations (C.15) determine six linear equations for the six unknown values $(f^{(0)}, f^{(1)}, f^{(2)})$, and $(b^{(0)}, b^{(1)}, b^{(2)})$.

The solution procedure is omitted here. It turns out that for fairly general signal distributions, the strategy profile constitutes a sequential equilibrium.

C.5 Secret Price Cutting

We investigate the story of secret price cutting in Stigler (1964), and examine Stigler's intuition that collusion between firms is difficult to sustain when deviations from collusion by price discounts for customers are private information. We look at a repeated game with private monitoring in which each stage is a Bertrand duopoly game, and show a nontrivial sequential equilibria where both firms set next stage prices according to current sales, and the firms follow opposite rules of action.

Two firms i, $j \in \{1,2\}$ are permanent players and produce for infinitely many periods. The marginal cost is normalized to 0. In each period, the firms are competing for one customer who will purchase with an exogenous probability r. The firms simultaneously set prices a_i, $a_j \in [0,1]$, which are unobservable to each other. If the customer decides to buy, she will choose any firm with probability one half when both firms charge the same price, and choose the lower price firm with probability $1-\varepsilon$, where $0<\varepsilon<1/2$ and ε is close to 0. Patronage of a firm is its private information, so a buying customer means no business for the other firm. However, when the firm fails to sell its goods, the customer either chooses not to buy or to buy from the other firm. This scene is often observed in flea markets and farmers' markets when two separately located vendors sell identical products.

The stage game has no pure-strategy Nash equilibrium, as undercutting the opponent is not always the best response. To see this, let χ denote the value $\varepsilon/(1-\varepsilon)$. When $a_j \in [0,\chi]$, firm i is better off by charging $a_i = 1$. Its payoff is maximized as a higher price compensates for the loss in the probability of a customer visit, $1 \cdot \varepsilon > a_j \cdot (1-\varepsilon)$.

The stage game does have a symmetric mixed-strategy Nash equilibrium. Let $P(\cdot)$ denote the probability distribution function for the strategy, and the support of its density function is contained in $[\chi, 1]$ because no player would charge a price lower than χ. Firm j is indifferent in charging $a_j \in [\chi, 1]$ if firm i acts following probability P. Firm j's payoff is

$$r \cdot a_j \cdot \Big[\varepsilon\big(P(a_j) - P(\chi)\big) + (1-\varepsilon)\big(P(1) - P(a_j)\big)\Big] \;=\; r \cdot a_j \cdot \Big[\varepsilon P(a_j) + (1-\varepsilon)\big(1 - P(a_j)\big)\Big],$$

where $P(\chi) = 0$, $P(1) = 1$. For the payoff to be independent of a_j, we solve for P

$$P(a) \;=\; \frac{1-\varepsilon-(\varepsilon/a)}{1-2\varepsilon}.$$

In this stage game equilibrium, each firm receives an expected payoff $r\varepsilon$.

Collusion is possible under perfect public monitoring, and each firm would charge a monopoly price $a = 1$ and receive an expected payoff $r/2$, which is greater than $r\varepsilon$. However, this is difficult under private monitoring because sales z_i, $z_j \in \{0, 1\}$ as a way of private monitoring are negatively correlated. Firm i's observation of $z_i = 1$ indicates $z_j = 0$, while $z_i = 0$ indicates $z_j = (1-r)\cdot 0 + r \cdot 1$, where $r > 0$. Note that Bayesian learning of a_j based upon current period history $a_i z_i$ is statistically independent of z_j's inference.

We now construct a nontrivial sequential equilibrium of the repeated Bertrand duopoly game, in which the monopoly price $a_i = a_j = 1$ can be sustained with positive probability. The profile requires that one firm, say i, acts as a leader and charges a monopoly price for the next period if there is a sale, while the other firm j, as a follower, does so only when it has no customer. In other cases, both firms charge a price $\tilde{a}$, which is chosen before the game starts and fixed so that intertemporal incentives are supported. Other prices are off the equilibrium paths. A different choice of $\tilde{a}$ may determine a different sequential equilibrium.

In the automaton representation of the strategy profile, the two-state state space is the same for both firms, $\mathcal{W} = \{w^R, w^P\}$, where w^R denotes the state of *collusion* and w^P denotes the state of *punishment*. The initial state can be a probability distribution over $\mathcal{W}$ but for simplicity we set it as w^R. The output function for both firms is the same

$$f(w^R) = 1, \quad \text{and} \quad f(w^P) = \tilde{a},$$

where $\tilde{a} \in [\chi, 1]$ is to be determined, but the transition functions are opposites,

$$\tau(w, a_i z_i) = \begin{cases} w^R, & \text{if } z_i = 1, \\ w^P, & \text{if } z_i = 0, \end{cases}$$

with

$$\tau(w, a_j z_j) = \begin{cases} w^R, & \text{if } z_j = 0, \\ w^P, & \text{if } z_j = 1, \end{cases}$$

where a_i, a_j, z_i, and z_j are the firms' current period prices and customer visits, respectively.

As usual, let β_i^t denote the probability that firm i's belief $\mathfrak{B}_i^t$ assigns to the case of firm j being in collusion state w^R when firm i experiences h_i^t, and likewise for β_j^t. The belief operators az^* are given as follows.

$$
\begin{aligned}
a_i z_i^*(\beta) &= 1, && \text{if } z_i = 1, \\
a_i z_i^*(\beta) &= 1 - r, && \text{if } z_i = 0, \\
a_j z_j^*(\beta) &= 0, && \text{if } z_j = 1, \\
a_j z_j^*(\beta) &= r, && \text{if } z_j = 0.
\end{aligned}
\tag{C.16}
$$

The belief sequences $\{\beta_i^t\}_{t=0}^{\infty}$ and $\{\beta_j^t\}_{t=0}^{\infty}$ are defined by $a_i z_i^*$ and $a_j z_j^*$, and each takes only two values for $t \geq 1$ regardless of the initial beliefs β_i^0 and β_j^0.

The value functions for leading firm i satisfy equations

$$
\begin{aligned}
V_i(w^R, \beta) = {} & (1-\delta) r \Big[\beta \frac{1}{2} + (1-\beta)\varepsilon\Big] \cdot 1 \\
& + \delta \Big[r\big(\beta \frac{1}{2} + (1-\beta)\varepsilon\big) V_i(w^R, 1) + \Big((1-r) + r\big(\beta \frac{1}{2} + (1-\beta)(1-\varepsilon)\big)\Big) V_i(w^P, 1-r) \Big]
\end{aligned}
\tag{C.17}
$$

by pricing at 1, and

$$
\begin{aligned}
V_i(w^P, \beta) = {} & (1-\delta) r \Big[\beta(1-\varepsilon) + (1-\beta)\frac{1}{2}\Big] \cdot \tilde{a} \\
& + \delta \Big[r\big(\beta(1-\varepsilon) + (1-\beta)\frac{1}{2}\big) V_i(w^R, 1) + \Big((1-r) + r\big(\beta\varepsilon + (1-\beta)\frac{1}{2}\big)\Big) V_i(w^P, 1-r) \Big]
\end{aligned}
$$

by pricing at $\tilde{a}$. Those for follower firm j satisfy

$$
\begin{aligned}
V_j(w^R, \beta) = {} & (1-\delta) r \Big[\beta \frac{1}{2} + (1-\beta)\varepsilon\Big] \cdot 1 \\
& + \delta \Big[r\big(\beta \frac{1}{2} + (1-\beta)\varepsilon\big) V_j(w^P, 0) + \Big((1-r) + r\big(\beta \frac{1}{2} + (1-\beta)(1-\varepsilon)\big)\Big) V_j(w^R, r) \Big]
\end{aligned}
\tag{C.18}
$$

by pricing at 1, and

$$V_j(w^P,\beta) = (1-\delta)r\Big[\beta(1-\varepsilon)+(1-\beta)\frac{1}{2}\Big]\cdot\tilde{a}$$
$$+\delta\Big[r\big(\beta(1-\varepsilon)+(1-\beta)\frac{1}{2}\big)V_j(w^P,0)+\Big((1-r)+r\big(\beta\varepsilon+(1-\beta)\frac{1}{2}\big)\Big)V_j(w^R,r)\Big]$$

by pricing at $\tilde{a}$.

To solve equations (C.17) and (C.18), we first set $\beta = 1,\ 1-r,\ r,\ 0$, and solve for $V_i(w^R,1)$, $V_i(w^P,1-r)$, $V_j(w^R,r)$, and $V_j(w^P,0)$ as functions of δ, r, ε, and $\tilde{a}$, and then use one of the equations in (C.17) or (C.18) to identify $V_i(w^R,\beta)$, $V_i(w^P,\beta)$, $V_j(w^R,\beta)$, and $V_j(w^P,\beta)$.

Here, we show a numerical example instead of the convoluted results of calculation. Let $\delta = 0.9$, $r = 0.8$, $\varepsilon = 0.1$, and $\tilde{a} = 0.4$. For each firm, the stage game Nash equilibrium payoff is $r\varepsilon = 0.08$, and the expected payoff from collusion is $r/2 = 0.4$. For leading firm i, the region of beliefs consistent with state w^R is $(0.7208, 1]$. The firm expects a payoff of $V_i(w^R,1) = 0.2905$ from collusion state w^R and $V_i(w^P,1-r) = 0.2703$ in punishing state w^P. For follower firm j, the region of beliefs consistent with state w^R is $(0.1365, 1]$. The firm expects a payoff of $V_j(w^R,r) = 0.2795$ from collusion state w^R and $V_j(w^P,0) = 0.2608$ in punishing state w^P. Efficiency is improved from the repeated play of the stage game Nash equilibrium strategy.

Of course, both firms can randomly decide which one is the leading player at the initial stage of the game, which is similar to the examples discussed by Phelan and Skrzypacz (2012).

D Proofs

Proof of Proposition 2. Given a belief-dependent behavior strategy profile $\mathfrak{s}$, we show that the claim holds true for an arbitrary player i. Consider beliefs $\mathfrak{B}_i$ over $\prod_{j\neq i}\Omega_j$. We construct a partition of $\mathfrak{B}_i$, $\mathcal{W}_i = \{w^{\alpha_i}\}_{\alpha_i\in\Delta(A_i)}$, such that w^{α_i} collects all beliefs that assign action choice $\alpha_i \in \Delta(A_i)$. For continuation of play, the states of beliefs of any of i's opponents, j, matter. Note that initial beliefs and finite-period histories uniquely identify the interim beliefs, so i's beliefs over her opponents' beliefs, $\mathcal{W}_j = \{w^{\alpha_j}\}_{\alpha_j\in\Delta(A_j)}$, summarize all i's information about the states of her opponents.

Abusing notations, let β_i denote probability distributions over $\prod_{j\neq i}\mathcal{W}_j$. By Proposition 1, $\mathcal{W}_i$, β_i are sufficient statistics for the continuation of play.

The belief $\beta_i^t \in w^{\alpha_i}$, an update of the initial beliefs for one of i's histories, is i's initial belief for period t. At the end of period t, Bayesian learning $\beta_i^{t+1} = a_i z_i^*(\beta_i^t)$, belongs to some w^{α_i}, where $a_i z_i^*$ is a belief operator, and $a_i \in A_i$ and $z_i \in Z_i$ are outcomes of $\mathfrak{s}_i$, because $\mathcal{W}_i$ is a partition. By the construction of $\mathcal{W}_i$, we define decision rule $f_i(w^{\alpha_i}) = \alpha_i$ and transition function $\tau(w, a_i, z_i) = w^{\alpha_i}$, if $a_i z_i^*(\beta) \in w^{\alpha_i}$ for $\beta \in w$. If in some period, $a_i z_i^*(\beta) \in w^{\alpha_i}$, while $a_i z_i^*(\beta') \in w^{\alpha_i'}$ for β, $\beta' \in w$, then further divide w into a collection of subsets. Because $|A_i|$ and $|Z_i|$ are finite, the process will stop within finite steps in any period.

We now have an automaton representation of the strategy profile $\mathfrak{s}$. □

Proof of Proposition 5. We let $\mathcal{W}_i = \{w_i^{(k)}\}_{k=1}^{\kappa_i}$, $\prod_{j\neq i}\mathcal{W}_j = \{\prod_{j\neq i}^{(k')} w_j\}_{k'=1}^{\eta_i}$, and let $v_{kk'}$ denote $v_{w_i^{(k)}}(\prod_{j\neq i}^{(k')} w_j)$. There are $\kappa_i \times \eta_i$ unknown parameters, $v_{kk'}$. If we can identify the parameters, then we are done.

In equation (2.1), the left hand side is $\beta \cdot v_{k\cdot}$, where $v_{k\cdot} = (v_{k1}, \dots, v_{k\eta_i})$, $k = 1, \dots, \kappa_i$. On the right, the term with $1-\delta$ is constant given $w_i^{(k)}$, f_i, f_j, and β. For the second term with δ, the list of a_i is determined by $w_i^{(k)}$ and f_i. If we express V_i on the right hand side in linear forms, the coefficients of $v_{\tilde{k}\tilde{k}'}$ are

$$Prob\big(z_i|f_i;\beta,f_j;\pi(z_i|a)\big)\cdot Prob\big(a_i z_i^*(\beta)|f_i;\beta,f_j;z_i\big) = Prob\big(a_i z_i^*(\beta), z_i|f_i;\beta,f_j;\pi(z_i|a)\big),$$

the joint distribution of signal z_i and all the states in $\mathcal{W}_j$ given by the Bayes' rule. Here, we abuse notations, e.g., $Prob(a_i z_i^*(\beta)|f_i;\beta,f_j;z_i)$ is the posterior probability of a state vector in $\prod_{j\neq i}\mathcal{W}_j$ conditioned on the action choices of player i, j, and signal z_i observed by player i.

Note that (2.1) holds true for arbitrary β. Although β has dimension $\eta_i - 1$, it has η_i degree of freedom. So for each given $w_i^{(k)}$, $k = 1, \dots, \kappa_i$, there are η_i linear equations of $v_{\tilde{k}\tilde{k}'}$, $\tilde{k} = 1, \dots, \kappa_i$, $\tilde{k}' = 1, \dots, \eta_i$. When k goes through all states in $\mathcal{W}_i$, we obtain $\kappa_i \times \eta_i$ linear equations in total. If the automaton representation of the strategy profile is minimal, then there exists a unique solution of $v_{kk'}$. Otherwise, there are a continuum of solutions. This finishes the proof. □

Proof of Proposition 4. Given a sequentially rational pure-strategy profile $\mathfrak{s}$, in its minimal automaton representation, $\{\mathcal{W}_i\}_{i=1}^n$ are finite sets. We continue to use the notations in the proof of Proposition 5.

The sequential rationality implies that the inequalities given by the one-shot deviation principle hold true for any β consistent with one of the states w. For any state $w_i^{(k)} \in \mathcal{W}_i$, $k = 1, \dots, \kappa_i$, it is a necessary condition that the inequalities still hold if we replace an arbitrary action a_i on the right hand side of the inequalities with $f_i(w_i^{(\bar{k})})$, $\bar{k} = 1, \dots, \kappa_i$, $\bar{k} \neq k$, and $f_i(w_i^{(\bar{k})}) \neq f_i(w_i^{(k)})$. By Proposition 5, the linearity in β of the value of continuation is true regardless of the value of β. The right hand side of the inequalities turn into $V_i(\tau(w_i^{(k)}, f_i(w_i^{(\bar{k})}), z_i), \beta)$, which are different from $V_i(w_i^{(k)}, \beta)$, after we replace a_i with $f_i(w_i^{(\bar{k})})$.

The inequalities are $\beta \cdot v_{k\cdot} \geq \beta \cdot v_{\bar{k}'\cdot}$. Here, we use $\bar{k}'$ to indicate that we may not obtain $V_i(w_i^{(\bar{k})}, \beta)$ if we replace a_i with $f_i(w_i^{(\bar{k})})$ on the right hand side. The only condition we are certain is that $\bar{k}' = 1, \dots, \kappa_i$, $\bar{k}' \neq k$, and $f_i(w_i^{(\bar{k}')}) \neq f_i(w_i^{(k)})$.

Without loss of generality, let $f_i(w_i^{(k)}) = a_i$. Although $w_i^{(k)}$ is a collection of infinitely many history paths, the paths can be grouped into finitely many categories because of the finite-state automaton representation. Given one of the categories and $f_i(w_i^{(\bar{k})}) = \bar{a}_i$, $\tau(w_i^{(k)}, f_i(w_i^{(\bar{k})}), z_i)$ must be some $\tau(w_i^{(\bar{k}')}, f_i(w_i^{(\bar{k}')}), z_i)$ because the number of the states is finite. Specifically, first, once the action $f_i(w_i^{(\bar{k})})$ is chosen, the probability distribution of signal z_i will be determined; second, the histories with one-period-lag in some category of $w_i^{(k)}$ concatenating with $f_i(w_i^{(\bar{k})})z_i$, where it should be $f_i(w_i^{(k)})z_i$, constitutes the right hand side in the value of continuation of $w_i^{(\bar{k}')}$ given by the one-shot deviation principle. Now we have $V_i(w_i^{(k)}, \beta) \geq V_i(w_i^{(\bar{k}')}, \beta)$, and there are usually more than one such inequalities as we go through all categories in $w_i^{(k)}$.

We go back to use $\bar{k}$ to denote the states different from $w_i^{(k)}$. For each inequality $\beta \cdot v_{k\cdot} \geq \beta \cdot v_{\bar{k}\cdot}$, it defines a half-space in $\mathbb{R}^n$, which contains the range of β. Each β^w, $w \in \mathcal{W}_i$, is contained in one of the intersections of these half-spaces and $\mathcal{M}_i$. Note that $\mathcal{M}_i$ and half-spaces are all convex and connected, and any intersection of such sets is also convex and connected. No two of these intersections are overlapped. At this point, it is straightforward to find a required partition $\mathcal{P}_i$ of $\mathcal{M}_i$ with each submanifold of $\mathcal{P}_i$ covering one of the intersections. This finishes the proof. $\square$

Proof of Proposition 7. Given $w^0 \in \Gamma(\mathcal{W})$, where w^0 is a collection of functions w_i^0, we construct a strategy profile $\mathfrak{s}$, such that it yields continuation promises of payoffs w^0 and $\mathfrak{s}$ is sequential. Then $\mathcal{V} \subset \mathcal{E}(\delta)$ follows. The idea is to build Bayesian stage games which are interconnected by outcomes $a_i z_i$, $a_i \in A_i$, $z_i \in Z_i$, and $i = 1, \ldots, n$.

In an automaton representation, states of player i are her last period outcomes $a_i z_i$ and state transition τ_i is trivial. The same outcome in a different period is treated as a different state, so the representation may not be minimal. To construct decision function f_i, we repeatedly use the assumption $\mathcal{W} \subset \Gamma(\mathcal{W})$. In period t, $t \geq 1$, there exists $w_i^{t-1} \in \mathcal{W}_i$ such that $\mathbf{U}_i\big(w_i^{t-1}\big)(a_i^t z_i^t) = w_i^t$, $w_i^t \in \mathcal{W}_i \subset \Gamma(\mathcal{W}_i)$ and $a_i^t z_i^{t*}\Big(\mathbf{B}_i\big(w_i^{t-1}\big)\Big) \subset \mathbf{B}_i\big(w_i^t\big)$. Then, $\mathbf{Q}_i\big(w_i^t\big) = \alpha_i^{t+1} \in \Delta\big(A_i\big)$, and we define $f_i(a_i^t z_i^t) = \alpha_i^{t+1}$. The automata are well-defined for all players on $\mathcal{W}$ because of the consistency of decomposition, but each automaton of itself differs in initial payoffs w_i^0 and the choices of w_i^{t-1}, w_i^t, and α_i^{t+1} for player i.

We need to show that, for each $w^0 \in \Gamma(\mathcal{W})$, the automata $\big(\{a_i z_i\}, \emptyset, f_i, \tau_i\big)$, $i = 1, \ldots, n$, describe a sequential equilibrium with payoff w^0. If we show, for each player i and any belief $\beta_i^0 \in \mathbf{B}_i\big(w_i^0\big)$, $w_i^0 \cdot \beta_i^0$ is the value of continuation play for player i, then the claim follows by the decomposability of w^0 and the one-shot deviation principle.

For any w_i^0 of payoff $w^0 \in \Gamma(\mathcal{W})$, we define an implied sequence of continuations $\{\gamma_i^t\}_{t=0}^{\infty}$ and $\{\varphi_i^t\}_{t=0}^{\infty}$, where $\gamma_i^t : A_i^t \times Z_i^t \to \mathcal{W}_i$ and $\varphi_i^t : \mathcal{B}_i^t \times A_i^t \times Z_i^t \to \mathcal{B}_i^{t+1}$. We set $\gamma_i^t = \mathbf{U}_i\big(w_i^t\big)$ and $\varphi_i^t = \mathbf{Y}_i\big(w_i^t\big)$. By the consistency of decomposition, $\gamma^t \equiv \{\gamma_i^t\}_{t=0}^{\infty}$ and $\varphi^t \equiv \{\varphi_i^t\}_{t=0}^{\infty}$ are consistent. Associated with each w_i^t, $\mathbf{B}_i\big(w_i^t\big)$ is its set of supporting beliefs. Let β_i^0 be an arbitrary element in $\mathbf{B}_i\big(w_i^0\big)$. Let $\mathfrak{s}$ be the strategy profile described by the automata $\big\{(\{a_i z_i\}, \emptyset, f_i, \tau_i)\big\}_{i=1}^{n}$, so that $\mathfrak{s}_i(\emptyset) = \mathbf{Q}_i(w_i^0)$, and for any history, h_i^t, ending in $a_i z_i$, $\mathfrak{s}_i(h_i^t) = f_i(a_i z_i)$.

By the construction and consistency,

$$
\begin{aligned}
w_i^0 \cdot \beta_i^0 &= \mathbf{V}_i\big(\mathbf{Q}_i(w_i^0), \beta_i^0, \mathbf{U}_i(w_i^0), \mathbf{Y}_i(w_i^0)\big) = \mathbf{V}_i\big(\mathfrak{s}_i(\emptyset), \beta_i^0, \mathbf{U}_i(w_i^0), \mathbf{Y}_i(w_i^0)\big) \\
&= (1-\delta)u_i\big(\mathfrak{s}(\emptyset)\big) + \delta \sum_{a_i^0} \sum_{z_i^0} \mathbb{E}^{\mathfrak{s}(\emptyset),\pi}\Big[\gamma_i^0(a_i^0, z_i^0) \cdot \varphi_i^0(\beta_i^0, a_i^0, z_i^0)\Big] \\
&= (1-\delta)u_i\big(\mathfrak{s}(\emptyset)\big) + \delta \sum_{h_i^1} \mathbb{E}^{\mathfrak{s}(\emptyset),\pi}\Big[w_i^1 \cdot \beta_i^1\Big] \\
&= (1-\delta)u_i\big(\mathfrak{s}(\emptyset)\big) + \delta \sum_{h_i^1} \mathbb{E}^{\mathfrak{s}(\emptyset),\pi}\Bigg[(1-\delta)u_i\big(\mathfrak{s}(h_i^1)\big) \\
&\qquad + \delta \sum_{a_i^1} \sum_{z_i^1} \mathbb{E}^{\mathfrak{s}(h_i^1),\pi}\Big[w_i^2 \cdot \beta_i^2\Big]\Bigg] \\
&= (1-\delta)\sum_{s=0}^{t-1} \delta^s \sum_{h_i^s} \mathbb{E}^{\mathfrak{s},\pi}\Big[u_i\big(\mathfrak{s}(h_i^s)\big)\Big] + \delta^t \sum_{a_i^{t-1}} \sum_{z_i^{t-1}} \mathbb{E}^{\mathfrak{s},\pi}\Big[w_i^t \cdot \beta_i^t\Big],
\end{aligned}
$$

where $\mathbb{E}^{\mathfrak{s},\pi}$ takes the probabilities of all possible history h_i^s under $\mathfrak{s}$. Because $w_i^t \in \mathcal{W}_i$, a bounded set, $\sum_{a_i^{t-1}} \sum_{z_i^{t-1}} \mathbb{E}^{\mathfrak{s},\pi}\big[w_i^t \cdot \beta_i^t\big]$ is bounded. Let $t \to \infty$, and we have

$$
w_i^0 \cdot \beta_i^0 = (1-\delta)\sum_{s=0}^{\infty} \delta^s \sum_{h_i^s} \mathbb{E}^{\mathfrak{s},\pi}\Big[u_i\big(\mathfrak{s}(h_i^s)\big)\Big],
$$

the value of $\mathfrak{s}$ with initial belief β_i^0. We have shown that, for each player i and all $w_i^0 \in \Gamma(\mathcal{W}_i)$ with automaton $\big(\{a_i z_i\}, \emptyset, f_i, \tau_i\big)$, $w_i^0 \cdot \beta_i^0$ is the payoff of continuation play.

We examine the value

$$
\begin{aligned}
& (1-\delta)u_i(\alpha) + \delta \sum_{a_i} \sum_{z_i} \mathbb{E}^{\alpha,\pi}\Big[\mathbf{U}_i(w_i)(a_i z_i) \cdot \mathbf{Y}_i(w_i)_{a_i z_i}(\beta_i)\Big] \\
= \quad & (1-\delta)u_i(\alpha) + \delta \sum_{a_i} \sum_{z_i} \mathbb{E}^{\alpha,\pi}\Big[\gamma_i(a_i, z_i) \cdot \varphi_i(\beta_i, a_i, z_i)\Big].
\end{aligned}
$$

Because w_i is consistently decomposed by $\big(\mathbf{Q}_i(w_i), \beta_i, \mathbf{U}_i(w_i), \mathbf{Y}_i(w_i)\big)$, and $\mathbf{Q}_i(w_i)$ is enforced by $\big(\mathbf{U}_i(v_i), \mathbf{Y}_i(w_i)\big)$ on $\mathcal{W}_i$ for i, $i = 1, \ldots, n$, the sequentiality follows by the one-shot deviation principle. □

Proof of Proposition 8. By Proposition 7, it suffices to show that $\mathcal{E}(\delta)$ is self-generating with respect to Γ.

Suppose that $\nu^0 \equiv \{\nu_i^0\}_{i=1}^n \in \mathcal{E}(\delta)$, and $\mathfrak{s} \equiv \{\mathfrak{s}_i\}_{i=1}^n$ is a sequential equilibrium such that player i expects the value of $\nu_i^0 = w_i^0 \cdot \beta_i^0$. The strategy profile $\mathfrak{s}$ has an automaton representation for each player and beliefs are consistent. For each player i, let $\alpha_i \equiv \mathfrak{s}_i(\emptyset)$ and $\gamma_i(a_i^0, z_i^0) \cdot \varphi_i(\beta_i^0, a_i^0, z_i^0) = u_i\big(\mathfrak{s}_i(h_i^1)\big)$, where $\varphi_i(\cdot, a_i^0, z_i^0)$ is defined by $\mathfrak{s}$ following the signal distributions.

Belief β_i^0 is the initial belief of the profile $\mathfrak{s}_i$. It is sufficient to show that α_i is enforced by γ_i and φ_i with initial belief β_i^0, and $\mathbf{V}_i(\alpha_i, \beta_i^0, \gamma_i, \varphi_i) = w_i^0 \cdot \beta_i^0$. To see this,

$$\begin{aligned}
\mathbf{V}_i(\alpha_i, \beta_i^0, \gamma_i, \varphi_i) &= (1-\delta)u_i(\alpha) + \delta \sum_{a_i} \sum_{z_i} \mathbb{E}^{\alpha,\pi}\Big[\gamma_i(a_i, z_i) \cdot \varphi_i(\beta_i^0, a_i, z_i)\Big] \\
&= (1-\delta)u_i(\alpha) + \delta \sum_{a_i} \sum_{z_i} \mathbb{E}^{\alpha,\pi}\Big[u_i\big(\mathfrak{s}_i(h_i^1)\big)\Big] \\
&= u_i(\mathfrak{s}_i) = \nu_i^0,
\end{aligned}$$

where α_{-i} is constructed from β_i^0. Because $\mathfrak{s}$ is sequential, $\{\mathfrak{s}_i(h_i^1)\}_{i=1}^n$ for any $h_i^1 = \{a_i^0 z_i^0\}$ is also sequential. Then, $\{\gamma_i(a_i, z_i)\}_{i=1}^n \in \mathcal{W}$ that calculates $\mathcal{E}(\delta)$. Moreover, for each player i, there is no profitable one-shot deviation, and α is enforced by γ_i and φ_i. We have $\nu^0 \in \Gamma\big(\mathcal{E}(\delta)\big)$. □

Proof of Lemma 3. Suppose $w = \{w_i\}_{i=1}^n \in \Gamma\big(\mathcal{W}\big)$. For each player i, $w_i \cdot \beta_i = \mathbf{V}_i\big(\alpha_i, \beta_i, \gamma_i, \varphi_i\big)$ for some α_i enforced by γ_i and φ_i on $\mathcal{W}_i$ with belief β_i. Note that belief inference φ_i remains the same and consistency still holds. Thus γ_i and φ_i decompose w_i using α_i on $\mathcal{W}_i'$, and we have $w = \{w_i\}_{i=1}^n \in \Gamma\big(\mathcal{W}'\big)$. □

Proof of Lemma 4. The sets $A_i \times Z_i$, $i = 1, \ldots, n$, are finite by assumption, so take subsequence if necessary, suppose that $\{w^k\}_k$ is a sequence in $\Gamma(\mathcal{W})$ converging to w. For each player i, let $(\alpha_i^k, \beta_i^k, \gamma_i^k, \varphi_i)$ denote the associated sequence of enforceable action profiles and enforcing continuations, while φ_i remains the same. We have $w_i^k \cdot \beta_i^k = \mathbf{V}_i\big(\alpha_i^k, \beta_i^k, \gamma_i^k, \varphi_i\big)$. Because $\mathcal{B}_i$, $\prod_{i=1}^n \Delta(A_i)$, and $\mathcal{W}_i^{A_i \times Z_i}$ are compact, their set-wise products are also compact. We can find a converging subsequence of $(\alpha_i^k, \beta_i^k, \gamma_i^k, \varphi_i)$ and without loss of generality, we assume that the sequence converges to $(\alpha_i, \beta_i, \gamma_i, \varphi_i)$. The action profile α_i is enforced by γ_i, φ_i over $\mathcal{W}_i$ with $w_i \cdot \beta_i = \mathbf{V}_i\big(\alpha_i, \beta_i, \gamma_i, \rho_i\big)$.

So, $w_i \in \Gamma(\mathcal{W}_i)$. $\Gamma\big(\mathcal{W}\big)$ is closed. $\square$

Proof of Proposition 9. We need to show that $\mathcal{W}^\dagger_\infty \subset \Gamma\big(\mathcal{W}^\dagger_\infty\big)$. We pick any $w = \{w_i\}_{i=1}^n \in \mathcal{W}^\dagger_\infty$, then $w \in \Gamma^m\big(\mathcal{W}^\dagger\big)$ for all m. There exists $(\alpha_i^m, \beta_i^m, \gamma_i^m, \varphi_i^m)$ such that $w_i \cdot \beta_i^m = \mathbf{V}_i\big(\alpha_i^m, \beta_i^m, \gamma_i^m, \varphi_i^m\big)$ and $\gamma_i^m(a_i, z_i) \in \Gamma^{m-1}(\mathcal{W}_i^\dagger)$ for all $a_i \in A_i$ and $z_i \in Z_i$. Note that β_i^m supports w_i for all m and Bayesian learning φ_i^m is determined once α_i^m, $i = 1, \ldots, n$, are chosen.

By compactness, and taking converging subsequence if necessary, we can assume that sequence $(\alpha_i^m, \beta_i^m, \gamma_i^m, \varphi_i^m)$ converges to a limit $(\alpha_i^*, \beta_i^*, \gamma_i^*, \varphi_i^*)$. We show that α_i^* is enforced by $(\gamma_i^*, \varphi_i^*)$ on $\mathcal{W}^\dagger_{i,\infty}$ with initial belief β_i^* and $w_i \cdot \beta_i^* = \mathbf{V}_i\big(\alpha_i^*, \beta_i^*, \gamma_i^*, \varphi_i^*\big)$. We need to verify $\gamma_i^*(a_i, z_i) \in \mathcal{W}^\dagger_{i,\infty}$, and the rest follows by the convergence and the smoothness of $\mathbf{V}_i$. Suppose that there are some a_i, z_i such that $\gamma_i^*(a_i, z_i) \notin \mathcal{W}^\dagger_{i,\infty}$. Because $\mathcal{W}^\dagger_{i,\infty}$ is closed, we can find an $\varepsilon > 0$ such that

$$\prod_{i=1}^n \bar{B}_\varepsilon\big(\gamma_i^*(a_i, z_i)\big) \cap \mathcal{W}^\dagger_\infty = \emptyset,$$

where $\bar{B}_\varepsilon(w_i)$ is a closed ball centered at w_i with radius ε in the functional space with supreme norm. But there exists m' such that for all $m > m'$, $\gamma_i^m(a_i, z_i) \in \bar{B}_\varepsilon\big(\gamma_i^*(a_i, z_i)\big)$, which implies

$$\bar{B}_\varepsilon\big(\gamma_i^*(a_i, z_i)\big) \cap \big(\cap_{m \le M} \Gamma^m(\mathcal{W}_i^\dagger)\big) \neq \emptyset, \quad \forall M > m',\ i = 1, \ldots, n.$$

And then the closed ball $\bar{B}_\varepsilon\big(\gamma_i^*(a_i, z_i)\big)$ has nonempty intersection with any finite collection of $\big\{\Gamma^m(\mathcal{W}_i^\dagger)\big\}_{m=1}^\infty$, and by the compactness of $\prod_{i=1}^n \bar{B}_\varepsilon\big(\gamma_i^*(a_i, z_i)\big) \cup \mathcal{W}^\dagger$,

$$\prod_{i=1}^n \bar{B}_\varepsilon\big(\gamma_i^*(a_i, z_i)\big) \cap \mathcal{W}^\dagger_\infty \neq \emptyset,$$

a contradiction. So $\mathcal{W}^\dagger_\infty$ is self-generating and also bounded, by Proposition 7, its associated $\mathcal{V}^\dagger_\infty \subset \mathcal{E}(\delta)$. But we know that $\mathcal{E}(\delta) \subset \mathcal{V}^\dagger_\infty$, and this finishes the proof. $\square$

References

Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), 1041–1063.

Amarante, M. (2003): "Recursive Structure and Equilibria in Games with Private Monitoring," *Economic Theory*, 22(2), 353–374.

Bhaskar, V., and I. Obara (2002): "Belief-Based Equilibria in the Repeated Prisoners' Dilemma with Private Monitoring," *Journal of Economic Theory*, 102(1), 40–69.

Bhaskar, V., and E. van Damme (2002): "Moral Hazard and Private Monitoring," *Journal of Economic Theory*, 102(1), 16–39.

Compte, O. (1998): "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, 66(3), 597–626.

Ely, J. C., J. Hörner, and W. Olszewski (2005): "Belief-free Equilibria in Repeated Games," *Econometrica*, 73(1), 377–415.

Ely, J. C., and J. Välimäki (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102(1), 84–105.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–554.

Iwasaki, A., Y. Joe, M. Kandori, I. Obara, and M. Yokoo (2012): "Automated Equilibrium Analysis of Repeated Games with Private Monitoring: a POMDP Approach," *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AA-MAS 2012)*, pp. 1305–1306.

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

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica*, 66(3), 627–652.

Kandori, M., and I. Obara (2006): "Efficiency in Repeated Games Revisited: The Role of Private Strategies," *Econometrica*, 74(2), 499–519.

——— (2010): "Towards a Belief-Based Theory of Repeated Games with Private Monitoring: An Application of POMDP," unpublished manuscript.

Mailath, G. J., and S. Morris (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102(1), 189–228.

——— (2006): "Coordination Failure in Repeated Games with Almost-Public Monitoring," *Theoretical Economics*, 1(3), 311–340.

Mailath, G. J., and L. Samuelson (2006): *Repearted Games and Reputations: Long-Run Relationships*. Oxford University Press, Inc., 198 Madison Avenue, New York, New York, 10016, 1st edn.

Phelan, C., and A. Skrzypacz (2012): "Beliefs and Private Monitoring," *The Review of Economic Studies*, 79(4), 1637–1660.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1), 70–83.

Sekiguchi, T. (1997): "Efficiency in Repeated Prisoner's Dilemma with Private Monitoring," *Journal of Economic Theory*, 76(2), 345–361.

Stigler, G. J. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 72(1), 44–61.