Essays on limited memory and information acquisition in long-run relationships

日野，喜文

(Degree)
博士（経営学）
(Date of Degree)
2015-09-25
(Date of Publication)
2016-09-01
(Resource Type)
doctoral thesis
(Report Number)
甲第6485号
(URL)
https://hdl.handle.net/20.500.14094/D1006485

博士論文

Essays on limited memory and information acquisition in long-run relationships

(長期的な関係における有限記録と情報収集に関する研究)

2015年7月21日提出

神戸大学大学院経営学研究科
宮原泰之研究室

経営学専攻
学籍番号 106B020B
氏名　日野喜文

Essays on limited memory and information acquisition in long-run relationships
(長期的な関係における有限記録と情報収集に関する研究)

氏名　日野喜文

Acknowledgments

I am deeply grateful to my adviser, Yasuyuki Miyahara, for his invaluable advice and guidance during every step of this dissertation. I am also grateful to my dissertation committee members, Hideo Suehiro, Hisao Hisamoto and Kenji Matsui for their encouragement and support.

Contents

Chapter 1

Introduction

In this dissertation, we analyze the nature of the long-run relationships that exist in infinitely repeated games. Many economic phenomena contain long-run relationships. We focus on two kinds of long-run relationships. One is the relationship between a long-run player and a sequence of short-run players. Consider a firm (a long-run player) and consumers (short-run players). A potential consumer who are considering to buy a product from a firm, acquires information about the quality of the product from past consumers. In such a situation, the past consumers' behavior affects future consumers' behavior and the firm's behavior. Thus, a firm and a sequence of consumers have a long-run relationship. The other is the relationship among multiple long-run players. Consider a team production in an organization. In most cases, the team produces products repeatedly with the same team members (long-run players). Thus, team members have long-run relationships.

In most studies that analyze long-run relationships in game theory, it is assumed that each player can store all information at no cost. We call these kinds of models *unlimited memory models*. However, unlimited memory models are unsuitable for studying some long-run relationships.

Imagine a restaurant and a (potential) consumer who is considering whether to go to the restaurant or not. Consumers acquire information about the restaurant from a guidebook. The consumer might not be able to acquire all the information about the restaurant from the guidebook. In general, a guidebook covers information about the current restaurant, however, it does not cover old information about the restaurant. In such case, old information about the restaurant is no longer available. In such situation, it may be plausible to assume that each player can store information only in the fixed number of previous periods. We call these kinds of models *limited memory models*. Limited memory models are more suitable for analyzing those long-run relationships than are unlimited memory models.

If information acquisition is costly, then players might not acquire information when he believes that he cannot obtain profitable information from costly information acquisition. Consider the following team production. A team member cannot observe other team members' behavior in a different division. However, if he incurs a cost and goes to the other team members' division, he can observe the behavior of the other team members. If the member believes strongly that other members choose a specific action profile, then he does not monitor the action profile and saves on the monitoring cost.

In this situation, it may be plausible to assume that each player can acquire the information if he pays the monitoring cost. We call these kinds of models *costly observation models*. These models are more suitable for studying long-run relationships than are unlimited memory models.

In this dissertation, we consider three models: unlimited memory models, limited memory models, and costly observation models. In what follows, we explain the results for each model.

First, we analyze unlimited memory models. We consider a *reputation model*, that is, a repeated game with a long-run player who has a type and a sequence of short-run players. Previous studies show the partial property of the equilibrium payoff set in reputation models. Fudenberg and Levine (1992) and Gossner (2011) consider infinitely repeated games with imperfect public monitoring. A monitoring structure is said to be *perfect* if each player observes the realized action profile in each period. A monitoring structure is said to be *imperfect public* if players cannot observe realized action profile, but they can observe the same noisy signals. They assume there exists a Stackelberg type as one of the commitment types. A Stackelberg type is a type who commits to play the action that the long-run player is most likely to commit to. They show that if there exists a Stackelberg type with a positive probability, then they find an upper bound and a lower bound of the equilibrium payoff set, which becomes tight as the discount factor goes to one. Cole and Kocherlakota (2001) present new methods to derive a set of pairs of equilibrium payoff vectors and common priors in a repeated game when each player can keep track of the other players' beliefs, and they find a sufficient condition that ensures that each player can keep track of the other players' beliefs.

We apply the idea of Cole and Kocherlakota (2001) to the case of reputation models that might not satisfy their sufficient condition. We show a method to derive the set of pairs of the equilibrium payoff vectors in which short run players use pure strategies and common priors in reputation models. We prove that, in reputation models, when short-run players use pure strategies, each player can keep track of the other players' beliefs on the equilibrium path, and we apply their technique. In addition, if the monitoring structure satisfies a certain standard assumption of reputation models (e.g., (a) the actions of short-run players are public, or (b) the probability distribution of public signals are independent of short-run players' actions), then we show that the methods of Cole and Kocherlakota (2001) hold in the case that players use mixed strategies.

The second model is a limited memory model. We assume that short-run players can observe a sequence of signals whose length is exogenously fixed. We consider a reputation model with a bad type, who commits to play the dominant action of the stage game. We consider infinitely repeated games in some class of games.

We focus on the following feature of reputation. Consider a firm and consumers.

When consumers believe that the firm produces a good product with similar probabilities, then they choose similar actions. Conjecturing such short-run players' behavior, the firm also chooses similar actions when consumers believe that the firm produces a good product with similar probabilities.

To capture the above feature of reputation, we focus on the following strategies as a class of equilibrium strategies: if a short-run player in a period has a similar belief to a short-run player in another period, then, similar mixed action profiles are chosen in these two periods.

In the limited memory model, a short-run player can observe the signal in the fixed number of previous periods, but cannot observe the signals in other periods. Hence, beliefs regarding the long-run player's type and his behavior change depending on the realized signals in those periods. Thus, the long-run player has a stronger incentive not to choose the dominant action of the stage game in a limited memory model than that in an unlimited memory model. Thanks to this strong incentive, we show that there exists an equilibrium in which the long-run player does not choose the dominant action of the stage game at any history in a limited memory model, although the long-run player chooses the dominant action of the stage game at any history in any equilibrium in an unlimited memory model.

The third model is a costly observation model. We consider a prisoner's dilemma as the stage game. We assume that a public randomization device is available, but communication is not available. Each player chooses whether to monitor the opponent player's action or not after his action choice. If he chooses to monitor, then he pays a cost and observes the opponent player's action; otherwise, he cannot obtain information about the opponent player's action at all. Hence, if a player pays a monitoring cost in each period, then he can keep track of the sequence of the opponent player's actions. Ben-Porath and Kahneman (2003) show the folk theorem with communication for arbitrary large monitoring costs. Miyagawa et al. (2003) show a sufficient condition of the folk theorem with communication for small monitoring costs. However, these two results do not cover an infinitely repeated prisoner's dilemma without communication. In this dissertation, we show an efficiency result when the observation costs are sufficiently small in repeated prisoner's dilemma. In general, to obtain efficiency result, many papers use complicated strategies (e.g., the six-state automaton strategy in Miyagawa et al. (2008)). However, the strategy we use in the proof of the efficiency result is less complex than that in previous works.

This dissertation is organized as follows. We provide an overview of the literature in Chapter 2. In Chapter 3, we study unlimited memory models. We study limited memory models in Chapter 4 and costly observation models in Chapter 5.

Chapter 2

Review of the Literature

2.1 Introduction

The following two features of long-run relationships enable agents to maintain good long-run relationships. One is that each agent can punish the other agent in the future when the other agent cheats on him. Hence, each player chooses a cooperative action to avoid the punishment. The other is that each agent can store information about the other agent. Sometimes, this generates an incentive for a player to choose a cooperative action in order to generate good information for him. Sometimes, it facilitates cooperation. Infinitely repeated games are suitable to analyze the former feature. Reputation models are suitable to analyze the latter feature. We review the literature on repeated games in Section 2.2 and reputation models in Section 2.3.

2.2 Repeated games

The monitoring structure of repeated games is important for punishing the other player in order to keep cooperation. Consider an infinitely played prisoner's dilemma. If each player observes the realized action profile (perfect monitoring), then players can maintain a good relationship using a grim trigger strategy. However, if players observe a signal that differs from the realized action profile with a positive error probability (imperfect public monitoring), then players no longer can maintain a good relationship using a grim trigger strategy. If each player observes different signals from each other (imperfect private monitoring), then players do not have common knowledge with the player they should punish. In such situations, it is not clear that players can keep cooperation. Many previous studies examine whether or not players can maintain good relationships under various monitoring structures.

2.2.1 Monitoring structure

The monitoring structure is *perfect* if each player observes the realized action profile. The monitoring structure is *imperfect public* if players cannot observe realized action profile, but they can observe the same noisy signals. The monitoring structure is said to be *imperfect private* if each player cannot observe the realized action profile, but he can observe different signals each other, which are realized stochastically and are his private information. The monitoring structure is costly observation if each player can observe the realized action profile when he pays a cost.

2.2.2 Public monitoring

Fudenberg and Maskin (1986) show that, in an infinitely repeated game with perfect monitoring, any individually rational payoff vector of a one-shot game of complete information can arise in an equilibrium if players are sufficiently patient (the so-called folk theorem). Some researchers try to extend their result to the case of imperfect public monitoring. Abreu et al. (1990) present an algorithm to check whether a set of payoffs is the equilibrium payoff set or not, with a fixed discount factor. From their algorithm, Fudenberg and Levine (1994) develop an algorithm to find the equilibrium payoff set when the discount factor goes to one. Fudenberg et al. (1994) present folk theorems in infinitely repeated games with imperfect public monitoring.

2.2.3 Private monitoring

Some studies analyze the cases of private monitoring. In this case, players do not have common knowledge about what the other players observe and it makes coordination difficult. A seminal work about infinitely repeated games with private monitoring has been conducted by Sekiguchi (1997). He shows that efficiency holds without communication nor public randomization in a repeated prisoner's dilemma game under private monitoring.

Many papers show the folk theorem with each private monitoring structure. Ben-Porath and Kahneman (1996), Kandori and Matsushima (1998), Compte (1998) and Obara (2009) present folk theorems in repeated games where communication is available. Ely and Välimäki (2002), Hörner and Olszewski (2006) show folk theorems without communication under almost perfect monitoring. Recently, Sugaya and Wolitzky (2014) present a folk theorem in a model with a mediator who can condition her recommendations on the entire history of actions and recommendations.

2.2.4 Costly observation

The above results assume that observation is costless. However, some papers assume that observation is costly. Lehrer (1989, 1992a, 1992b) considers two-player repeated games with no discounting. He shows that costly observation is supported in the equilibrium. Furthermore, he shows that an action profile can be played in the equilibrium even when there exists another short-run best reply that does not affect the other player's signal, if the player can obtain more precise information from the action than the short-run best reply. That is, he shows that costly observation is supported in the equilibrium.

Ben-Porath and Kahneman (2003) consider infinitely repeated games in which he observes the other players' actions if a player pays a cost, and, he does not observe at all otherwise. They assume that communication is available and show folk theorem even if the monitoring cost is high. Miyagawa et al. (2003) consider similar models to those in Ben-Porath and Kahneman (2003). Miyagawa et al. (2003) assume that communication is not available and each player can choose at least three kinds of actions. Miyagawa et al. (2003) show a sufficient condition for the folk theorem when the monitoring cost is small.

We consider an infinitely repeated prisoner's dilemma without communication. The results in Ben-Porath and Kahneman (2003) and Miyagawa et al. (2003) do not cover this game because communication is not available and each player can choose only two kinds of actions. We show that efficiency holds if a public randomization device is available.

The above two studies assume that if a player does not pay any monitoring cost, then he cannot observe at all. Miyagawa et al. (2008) consider a deviation from this assumption. They assume that each player can observe a private signal even if he does not incur a cost. Miyagawa et al. (2008) show a folk theorem in repeated games without communication for any level of monitoring costs. Miyagawa et al. (2008) use complicated strategies (e.g., six-state automata strategy in prisoner's dilemma) to prove their results. We focus on a three-state automaton strategy and show that efficiency holds when public randomization is available and the monitoring cost is sufficiently small.

Some studies assume that if a player incurs a cost, then he can obtain additional information but not about the action chosen by the other player in the current period. Kandori and Obara (2004) assume that the monitoring activity reveals not only the rival's action but also the "monitoring activity" and demonstrates efficiency in a certain class of strategies. Flesch and Perea (2009) assume that each player can additionally observe "the action chosen in the past" if he incurs additional cost. They show that if players can choose at least four actions, then the folk theorem holds even when neither public randomization nor communication is available. Awaya (2014) conducts a robustness check of the study of Takahashi (2010). Takahashi (2010) presents a folk theorem in repeated games where a continuum of players are randomly matched in each period to play the prisoner's dilemma with a different partner. Awaya (2014) introduces monitoring costs to the model of Takahashi (2010) and consider a model in which each player can observe "the sequence of actions that the opponent player chose in the past" if he observes the other player. He shows a folk theorem by constructing an equilibrium strategy when the monitoring cost is infinitesimal. That is, the monitoring cost is zero, but each player prefers not monitoring when monitoring decision does not affect his expected payoff (lexicographic preference). In contrast to the infinitesimal cost, he also shows that if the monitoring cost is strictly positive, then the strategy is not an equilibrium. In addition, if the monitoring cost is greater than the maximum difference of stage-game payoffs, then any equilibrium is a repetition of stage-game Nash equilibrium action profile. In Chapter 5 of the current dissertation, we assume that if a player incurs a cost, then he observes the action chosen by the other player, but he cannot obtain any information if he does not incur a cost. We show that efficiency holds when monitoring cost is sufficiently small and public randomization is available.

2.3 Reputation models

Another feature of long-run relationship is that each agent can store the information about the other agent. Hence, a long-run player has an incentive to choose an action in order to produce a good information for him. This feature dramatically changes long-run relationship. Especially, it changes long-run relationships when there exists incomplete information.

Let us consider reputation models. A reputation model is the following infinitely repeated game. There exist a long run player who has a type, and a sequence of short run players. In each period, a long-run player and a single short-run player play a stage-game. After players play a stage-game, the short-run player exits the infinitely repeated game, and a new short-run player participates in the infinitely repeated game and he plays a stage-game in the next period.

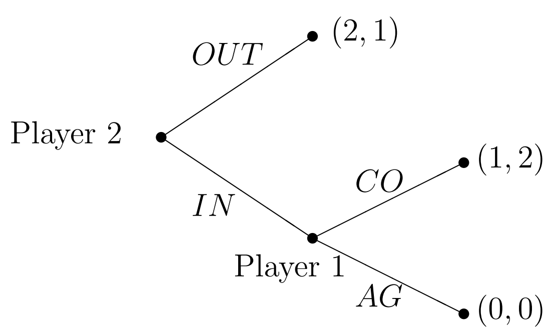
In reputation models, the realized signals could contain information about the type. Hence, if the long-run player is sufficiently patient, then he has an incentive to exert effort in order to give short-run players information that is favorable for him. This incentive derives from incomplete information about the characteristics. Thus, in reputation models, equilibrium behaviors differ substantially from those in repeated games with no incomplete information. In some models, this change excludes some equilibrium payoffs from the equilibrium payoff set.

We analyze reputation models with a bad type, who commits to playing the dominant action of the stage game. Hence, in particular, we review the literature about reputation models with bad types.

2.3.1 Standard reputation models

Selten (1978) considers the following *chain store game*. The chain store game has an incumbent (player 1) and an entrant (player 2). At the beginning of the chain store game, player 2 must decide either to enter the market (IN) or to stay out of the market (OUT). No further decisions are made if player 2's decision was OUT. If player 2's decision was IN, player 1 has to choose between cooperative behavior (CO) and aggressive behavior (AG). The payoff is summarized as follows.

Chain store games



Selten (1978) shows that player 1 cannot deter player 2's IN in the unique subgame perfect equilibrium of the finitely repeated chain store game. However, this result does not seem to be intuitive.

Kreps and Wilson (1982) and Milgrom and Roberts (1982) reexamine the chain store game in Selten (1978). They add to it incomplete information about player 1's type and illustrate the role of reputation. They consider a "tough" type, who chooses AG in each period, as one of player 1's possible types. If player 2 believes that player 1 is a tough type with high probability, then player 2 prefers OUT to IN in order to avoid (IN,AG). Hence, player 1 of "normal" type, who can choose his action between AG and CO in each period, has an incentive to choose AG to pretend to be the tough type in order to deter IN. They show that player 1 can succeed in deterring IN by playing AG in early periods and therefore obtains a higher equilibrium payoff than that in the chain store game in Selten (1978).

Repeated normal form games are often analyzed to study reputation models. In particular, the following product choice game is often analyzed.

Product choice game

Player 2

		h	ℓ
Player 1	H	$2, 3$	$0, 2$
	L	$3, -1$	$1, 1$

Action L is a strictly dominant action for player 1. Player 2's best reply is action h (resp. ℓ) if player 1 chooses action H (resp. L). In the following, we call an action a Stackelberg action if the long-run player is most likely to commit himself to play the action. We call a commitment type a *Stackelberg type* if the type commits to play a Stackelberg action in each period.

Fudenberg and Levine (1989) consider infinitely repeated normal form games. They generalize the results of Kreps and Wilson (1982) and Milgrom and Roberts (1982) to infinitely repeated games with perfect monitoring and show that if there exists a Stackelberg type with positive probability, then the infimum of the equilibrium set of a long-run player increases as the discount factor goes to one. Fudenberg and Levine (1992) and Gossner (2011) extend the results of Fudenberg and Levine (1989) to infinitely repeated games with imperfect public monitoring. In the literature on reputation models, such a change in the equilibrium payoff is referred to as a *reputation effect*.

2.3.2 Reputation of bad types

Bad type is a commitment type, who commits to play the dominant action of the stage game in each period. We can interpret a bad type as an inept type as follows. Consider a product choice game. A bad type is a player who has no ability to choose action H. In reality, there is a difference in the technologies of the firms. Some firms might have no ability to produce high-quality products. In that case, we interpret the firm as a bad type. To analyze such a situation, the study of the reputation model with a bad type is important for understanding the long-run relationship.

Recently, some studies introduced a bad type instead of a Stackelberg type. Bad types also have a dramatic effect on the equilibrium behavior, even if they offer very close to complete information.

Ely and Välimäki (2003) show that if there exists a bad type, then a supremum of the equilibrium payoff set in some reputation models converges to a stage-game Nash equilibrium payoff as the discount factor goes to one. That is, players fail to maintain good relationships. Ely et al. (2008) generalize Ely and Välimäki (2003) and consider games such that short-run players have an option of whether to deal with the long-run player who can choose a *deceiving action* to look good in the current period. Ely et al. (2008) call the generalized games *bad reputation games*. Ely et al. (2008) show that the supremum of the equilibrium payoff set converges to a stage-game Nash equilibrium payoff as the discount factor goes to one.

Our model in Chapter 3 is not a bad reputation game because the stage game in the dissertation has no deceiving action. We show that a unique sequential equilibrium in a class of equilibria in an unlimited memory model is a type of stage-game Nash equilibrium. Hence, our results are similar to those of Ely and Välimäki (2003) and Ely et al. (2008).

In contrast to Ely and Välimäki (2003) and Ely et al. (2008), Mailath and Samuelson (2001) show that a bad type helps to maintain a good reputation in some models. They consider a product choice game and *high-effort strategy*. A high-effort strategy is a strategy profile where player 1 chooses action H for any history, and short-run players choose best responses to it. If there exists no bad type, player 1 does not have an incentive to play action H because whatever signal is realized, the continuation strategy of player 2 is a repetition of action h. If there exists a bad type, player 1 has an incentive to play action H because player 2 changes her action to action ℓ when player 2 strongly believes that player 1 is a bad type.

However, even if there exists a bad type, a high-effort strategy is not a Nash equilibrium in an unlimited memory model. This is because player 1 loses the incentive to play action H at a point in time such that player 2 strongly believes that player 1 is a bad type. Mailath and Samuelson (2001) introduce stochastic change of player 1's type, that is, player 1's type changes in each period with positive probability.

Because of this stochastic change of player 1's type, old public signals become less informative about the present long-run player's type as time goes by. Hence, whatever signals are realized in the past, short-run players do not strongly believe that player 1 is a bad type. Thus, Mailath and Samuelson (2001) show that a high-effort strategy is a Nash equilibrium in reputation models with a bad type and stochastic change of player 1's type. That is, they show that a bad type and stochastic change of player 1's type lead to a cooperative relationship.

Our results in limited memory models in Chapter 4 are similar to those in Mailath and Samuelson (2001). We also find that a high-effort strategy is a Nash equilibrium in limited memory models. That is, we show that bad type and limited memory lead to a cooperative relationship. Thus, we find that limited memory and stochastic change of player 1's type have similar effects on reputation.

2.3.3 Limited memory

Recently, some works have tried to study reputation models under limited memory. Limited memory is a new model where short-run players keep track of the information from a fixed number of previous periods, but cannot keep track of the information in the other periods. Limited memory models are one of the important models for understanding long-run relationships. Consider a restaurant and consumers. Consumers buy a magazine about restaurants in their city to decide which restaurant they will choose. However, the magazine contains only recent information about the restaurant. Old information about the restaurant appears in an old magazine and it is no longer available. Hence, sometimes short-run players have limited memory. Analyses of reputation models with limited memory are important for understanding these long-run relationships.

Ekmekci (2011) considers a reputation model with a bad type. In his model, there exists a central mechanism that constructs a report from the sequence of signals in the past and sends it to short-run players. Short-run players receive the report, but cannot observe either signals or reports made in the past. Thus, short-run players have a one-period memory. He shows that there exists a central mechanism under which players can maintain a cooperative relationship in the long run. Our results in Chapter 4 also show that there exists a sequential equilibrium where players can maintain a cooperative relationship in the long run. However, our results do not need a central mechanism.

In the present dissertation, we define a limited memory model as a reputation model in which short-run players can observe signals in a fixed number of previous periods. Liu and Skrzypacz (2014) consider limited memory models with perfect monitoring. They show a unique perfect Bayesian equilibrium in which players play Pareto-efficient action profiles. Liu (2011) considers a reputation model with costly observation. He shows a unique perfect Bayesian equilibrium in which short-run players use random monitoring and players can maintain a good relationship in the long run. We find a pure strategy sequential equilibrium in which player 1 chooses action H for any history, although Liu and Skrzypacz (2014) and Liu (2011) find a mixed strategy equilibrium in which player 1 chooses action L with probability 1 at some history.

Liu and Skrzypacz (2014) and Liu (2011) assume that no calendar time is available to make their analyses simple. However, we and Monte (2013) assume that calendar time exists. Monte (2013) considers reputation models with various commitment types and shows that player 2 does not believe that player 1 is a specific type with high probability for any history. This property of limited memory helps to maintain good relationships. We show that there exists a sequential equilibrium in which a Pareto-efficient action profile is played for any history.

We show two new results in the study of reputation with limited memory. We focus on the following strategies as equilibrium strategies : if a short-run player in a period has a similar belief to a belief of a short-run player in another period, then similar mixed action profiles are chosen in those two periods. Ely et al. (2008) show that it converges to a stage-game Nash equilibrium payoff in a bad reputation game. However, we show that the supremum of the equilibrium payoff set does not converge to a stage-game Nash equilibrium payoff even when the discount factor goes to one in a limited memory model with a bad type. This result does not need either replacement of the long-run player's type as in Mailath and Samuelson (2001) or an exogenous central mechanism as in Ekmekci (2011). The other result is that any equilibria is a repetition of the stage-game Nash equilibrium action profile. Hence, we find that a limited memory helps players build and maintain a cooperative relationship.

2.4 Summary

We show the following results.

In Chapter 3, we analyze unlimited memory models. We apply the idea of Cole and Kocherlakota (2001) to reputation models in which short-run players use pure strategies. That is, we find an algorithm that solves for the set of pairs of the equilibrium payoff vectors in which short-run players use pure strategies and common priors in reputation models. In addition, if the standard assumption of reputation models (e.g., observability of short-run players' actions) holds, then we can apply the result of Cole and Kocherlakota (2001) to reputation models in which players use mixed strategies. These findings are new results about the equilibrium payoff set in reputation models.

In Chapter 4, we analyze a limited memory model. We consider infinitely repeated games with a bad type. Our study is the first attempt to study a combination of a limited memory model and a bad type. We focus on the following strategies as equilibrium strategies : if a short-run player in a period has a similar belief to a belief of a short-run player in another period, then similar mixed action profiles are chosen in those two periods. We show that a unique equilibrium in an unlimited memory model is a repetition of a stage-game Nash equilibrium. However, there exists an equilibrium in which the long-run player does not choose a dominant action of the stage game to build and maintain his reputation in a limited memory model. Based on these results, we find that limited memory leads to cooperative relationships.

In Chapter 5, we analyze costly observation models. We consider an infinitely repeated prisoner's dilemma without communication. The results in previous works about costly observation do not cover this game because communication is not available and players can choose only two kinds of actions. We show that if public randomization is available, then efficiency holds when the monitoring costs are sufficiently small.

Chapter 3

Equilibrium payoff set in reputation model

3.1 Introduction

In some models, incomplete information changes a long-run relationship dramatically. For example, consider an entry deterrence in a finitely repeated game. Selten (1978) shows that, in a finitely repeated chain store game, each entrant enters the market in the unique subgame perfect Nash equilibrium. That is, he shows that we cannot explain entry deterrence by finitely repeated games with complete information. However, if there exists incomplete information, entry deterrence can be explained by finitely repeated games. In fact, Kreps and Wilson (1982) and Milgrom and Roberts (1982) show that if there exists incomplete information about an incumbent's type in a finitely repeated chain store game, then entrants do not enter the market in the early periods on the sequential equilibrium path.

Repeated games under incomplete information is important, but it is more complex than games under complete information. Many papers show an upper bound and a lower bound of the equilibrium payoff set instead of characterizing the equilibrium payoff set. In this chapter, we consider reputation models, in which there exist a long-run player who has a type and a sequence of short -run players and analyze the following set of pairs of payoff vectors and common priors. We focus on the set of pairs (v, μ), where v is a vector of Nash equilibrium payoffs in which short-run players use pure strategies and μ is a common prior over the set of types. We show the following two results: (i) the set is the largest fixed point of a set-valued operator of the set of pairs of payoff vectors and probability distributions over the set of types, and (ii) the set is the limit of iterating this operator on any superset of it.

A large literature examines infinitely repeated games under incomplete information. The literature can be grouped roughly as follows. One group involves repeated games with long-run players. In most of the studies in this group, players are uncertain initially about the distribution of signals and their stage-game payoffs. Many studies (e.g., Wiseman (2005), Fudenberg and Yamamoto (2011) and Yamamoto (2013)) develop folk theorems under the condition that each player can statistically distinguish between any two states by realized signals (statewise full rank condition).

The other group analyzes repeated games in which there are two kinds of players: a long-run player and a sequence of short-run players who are uncertain about the types of the long-run player, which is the model we adopt in this chapter. Many studies (e.g., Fudenberg and Levine (1989, 1992) and Gossner (2011)) analyze the equilibrium payoff sets when the discount factor goes to one, and show that in some infinitely repeated games, the equilibrium payoff sets under incomplete information differ markedly from that under complete information.

Cole and Kocherlakota (2001) apply the results of Abreu et al. (1990) to the case of the infinitely repeated incomplete information games with multiple long-run players under some conditions. Abreu et al. (1990) present methods to derive the equilibrium payoff set in repeated games with imperfect public monitoring. A monitoring structure is said to be imperfect public if each player cannot observe realized action profile, but can observe the same signal, which is realized stochastically. It means that they cannot punish any player depending on actions. They show that if the full support assumption is satisfied in a infinitely repeated game with complete information, it holds that (i) the Nash equilibrium payoff set is the largest fixed point of a set-valued operator of the set of payoffs, and (ii) the Nash equilibrium payoff set is the limit of iterating this operator on any superset of the Nash equilibrium payoff set.

Cole and Kocherlakota (2001) analyze games in which the state of each player changes stochastically in each period. Their idea is as follows. If each player can keep track of the posterior beliefs of the other players, then any subgame can be analyzed as the game from the initial period given the posterior beliefs. They apply the idea of Abreu et al. (1990) and they present a method to derive a set of pairs of the equilibrium payoff vectors and common priors.

To ensure that each player can keep track of the posterior belief of the other players, they focus on *Markov-private strategies* and games in which beliefs are *Markov*. A strategy profile is *Markov-private* if each player's strategy depends on his past private information only through his current private state. A strategy profile is said to be Markov-private equilibrium if it is a sequential equilibrium and Markov-private. Given a game, beliefs are said to be *Markov* if, for any Markov-private strategy profile, each player's belief about the other players' states depend on his private information only through his current state, both on and off the equilibrium path.

They show the following two results about Markov-private equilibria. First, they define the set of pairs $(\tilde{v}, \tilde{\mu})$, where $\tilde{v}$ is a vector of Markov-private equilibrium payoffs and $\tilde{\mu}$ is the initial common prior over the set of states. They show that this set is the largest fixed point of a set-valued operator of the set of pairs of payoff vectors and probability distributions over the set of states. Second, they show that the largest fixed point is the limit of iterating this operator on a sufficiently large initial set of (v, μ) pairs.

Cole and Kocherlakota (2001) focus on games in which beliefs are Markov. However, beliefs are not Markov in many games. For example, consider the following team production game in which the probability of success depends on both players' effort choices. Each player can observe the output of the team production, but he cannot observe the effort choice of the other player. Nature selects players' types at the beginning of the repeated game. The state space is the type space of each player. Fix a carrot and stick strategy in which players choose high effort if and only if production occurs successfully in the previous period. Each player conjectures the types of the other member. This conjecture depends on not only his type (state) but also the action chosen by himself (his *private* information which is irrelevant to his type off the equilibrium path). Based on the same reasons as discussed above, belief are not Markov in many games.

The contribution of this chapter is to apply the idea of Cole and Kocherlakota (2001) to the case of the pure strategy equilibrium payoff sets in reputation models in which beliefs might not be *Markov*. We focus on reputation models which is more restrictive than their models in which they allow that states change throughout a play. We define the set of pairs of pure strategy equilibrium payoff v and common prior over the long-run player's type μ in the same way as Cole and Kocherlakota (2001). We show that (i) the set is the largest fixed point of a set-valued operator of the set of pairs of payoffs and probability distributions over the set of states, and (ii) the set is the limit of iterating this operator on any superset of it. In addition, we show that, these two results hold even when players can use mixed strategies under a class of standard reputation model assumptions; that is, (a) the distribution of public signals is independent of the short players' actions, or (b) the short-run players' actions are observable.

The key idea of this chapter is that if short-run players use pure strategies, then each player can keep track of the posterior beliefs of the other players on the equilibrium path in the same way as Cole and Kocherlakota (2001). Fix a pure strategy Nash equilibrium. Then, the long-run player can know what a short-run player plays in period 0 because short-run players use "pure" strategy. Hence, the long-run player knows short run player's private history in period 1 on the equilibrium path. Thus, the long-run player can know what a short-run player play in period 1 because short-run players use pure strategy. We can use the same argument in period $t = 0, 1, 2, \ldots$ Thus, a long run player precisely infers short run players' private history given any public history on the equilibrium path. Therefore, a long run player can keep track of short run players' belief on the equilibrium path. It implies that the idea of Cole and Kocherlakota (2001) can be applied to the case of reputation models in which beliefs might not be Markov.

In Section 3.6, we discuss sufficient conditions to apply the result of Cole and Kocherlakota (2001) to reputation models in which short-run players use mixed strategies.

The rest of this chapter is organized as follows. We introduce the model in Section 3.2. Section 3.3 studies the relationship between the Nash equilibrium and public equilibrium (PE) and Section 3.4 studies the public equilibria. Section 3.5 presents an example. Section 3.6 discusses the case in which players use mixed strategies. Section 3.7 concludes.

3.2 Model

We consider infinitely repeated games with imperfect public monitoring, and focus on the set of Nash equilibrium payoffs in which the short-run players' strategies are pure.

First, we explain an infinitely repeated game with complete information. Next, we introduce incomplete information to the infinitely repeated game.

3.2.1 The complete information game

The stage-game consists of player 1 and player 2. Let A_i be the finite set of player i's actions. Each player i chooses action $a_i \in A_i$. Let $u_i(a)$ be an expected stage-game payoff for player i given action profile $a \in A \equiv A_1 \times A_2$. Each player cannot observe the opponent player's action. Given action profile $a \in A$, a signal y is realized with probability $\rho(y|a)$ and each player observes the signal. We call y a public signal and we let Y denote the finite set of public signals.

We consider imperfect public monitoring. A monitoring structure is said to be *imperfect public* if players cannot observe realized action profile, but they can observe the same noisy signals. The following assumption ensures that, for any action profile, each player cannot observe the other player's action directly.

Assumption 3.1 (full support) For any $a \in A$, it holds that $\rho(y|a) > 0$ for any $y \in Y$.

By Assumption 3.1, for any Nash equilibrium, any public history is on the equilibrium path. Thus, each player cannot observe other player's deviation.

For any set Z, let us define $\Delta(Z)$ as the set of probability distributions over Z. Abusing notation, for any mixed action $\alpha_1 \in \Delta(A_1)$ and pure action $a_2 \in A_2$, we define $\rho(y|\alpha_1, a_2)$ and $u_i(\alpha_1, a_2)$, as follows.

$$\rho(y|\alpha_1, a_2) \equiv \sum_{a_1 \in A_1} \rho(y|a_1, a_2)\alpha_1(a_1), \quad \text{and}$$
$$u_i(\alpha_1, a_2) \equiv \sum_{a_1 \in A_1} u_i(a_1, a_2)\alpha_1(a_1).$$

We consider infinitely repeated games with imperfect public monitoring. Player 1 is a long-run player with discount factor $\delta \in [0, 1)$ and player 2 is a sequence of short-run players. At the beginning of each period $t = 0, 1, \ldots$, a signal x^t that is uniformly distributed over $[0, 1]$ is realized and each player observes the signal.

The timing in period t is summarized as follows. First, a sunspot x^t is realized. Second, each player i simultaneously chooses an action $a_i^t \in A_i$. Third, a public signal y^t is realized.

We denote by $h_i^t = (x^s, a_i^s, y^s)_{s=0}^{t-1}$, the private history of player i. We call a sequence of public signals in the past $(x^s, y^s)_{s=0}^{t-1}$ the public history and denote it by h^t. In what follows, let $\mathbb{N}$ be the set of natural numbers. For any $t \in \mathbb{N}$, $\mathcal{H}_i^t \equiv ([0, 1] \times A_i \times Y)^t$ is the set of private histories of player i in period t, and $\mathcal{H}^t \equiv ([0, 1] \times Y)^t$ is the set of public histories in period t. Let $\mathcal{H}_1^0, \mathcal{H}_2^0$ and $\mathcal{H}^0$ be any singleton set. That is,

$\mathcal{H}_1^0 = \mathcal{H}_2^0 = \mathcal{H}^0 = \{\phi\}$. Let us define $\mathcal{H}_i$ as the set of private histories of player i and $\mathcal{H}$ as the set of public histories. That is, $\mathcal{H}_i = \cup_{t=0}^{\infty} \mathcal{H}_i^t$ and $\mathcal{H} = \cup_{t=0}^{\infty} \mathcal{H}^t$. Player i's strategy is a function of $\mathcal{H}_i \times [0,1]$ to $\Delta(A_i)$. For expositional simplicity, if player i's strategy σ_i is pure, we regard σ_i as a function of $\mathcal{H}_i^t \times [0,1]$ to A_i instead of $\Delta(A_i)$.

3.2.2 The incomplete information game

Nature selects player 1's type from a finite set Ω according to common prior $\mu \in \Delta(\Omega)$. Player 1's type is his private information and does not change throughout a play. Let $u_{1,\omega}$ be the expected stage-game payoff function for player 1 of type ω. Let us denote by $\sigma_{1,\omega}$ the strategy of player 1 of type ω and denote $\sigma_1 = (\sigma_{1,\omega})_{\omega \in \Omega}$. Player 1 of type ω maximizes his discounted average expected stage-game payoffs given σ_2.

$$U_{1,\omega}(\sigma_{1,\omega}, \sigma_2) = (1-\delta) \sum_{t=0}^{\infty} \delta^t E[u_{1,\omega}(a^t) | \sigma_{1,\omega}, \sigma_2]$$

Let $\eta_2(\omega | h_2^t)$ be player 2's belief about the type of player 1 given the private history h_2^t, the common prior μ, and a strategy profile $\sigma \equiv (\sigma_1, \sigma_2)$ where we suppress μ, σ in η_2. Player 2 in period t maximizes the expected stage-game payoff in period t given σ_1, h_2^t and η_2. We use Nash equilibrium as a solution concept. Hence, we do not have to define beliefs over private histories.

Our model allows randomized commitment type, who play a mixed action profile in each period. Consider a type who are indifferent to any action profile and he chooses the same mixed action in each period. Then, we interpret this type as a randomized commitment type.

3.3 Nash equilibrium and public equilibrium

In this section, we show the key ideas of this chapter. First, we define *public equilibrium* (PE). Second, we show that, for any Nash equilibrium in which σ_2 is pure, there exists a PE in which each player obtains the same equilibrium payoff as that of the Nash equilibrium.

Definition 3.1 Player i's strategy σ_i is *public* if, for any $t \geq 1, h_i^t = (x^s, a_i^s, y^s)_{s=0}^{t-1}$, $\hat{h}_i^t = (\hat{x}^s, \hat{a}_i^s, \hat{y}^s)_{s=0}^{t-1}$ and $x \in [0,1]$, it holds that if $x^s = \hat{x}^s$ and $y^s = \hat{y}^s$ for any $s \leq t-1$, then $\sigma_i(h_i^t, x) = \sigma_i(\hat{h}_i^t, x)$.

Given common prior $\mu \in \Delta(\Omega)$, a strategy profile σ is *public equilibrium* (PE) if σ_1 and σ_2 are public and σ is a Nash equilibrium.

In what follows, for expositional simplicity, we denote the public strategy of player 1 of type ω, $\sigma_{1,\omega}$, by a function of $\mathcal{H} \times [0,1] \to \Delta(A_1)$ and the strategy of player 2 which is pure and public by a function of $\mathcal{H} \times [0,1] \to A_2$. We show that for any Nash equilibrium σ in which σ_2 is pure, there exists a PE in which each player obtains the same payoff as that of the Nash equilibrium.

Proposition 3.1 *Suppose that Assumption 3.1 is satisfied. For any $\mu \in \Delta(\Omega)$, fix a Nash equilibrium $\sigma = (\sigma_1, \sigma_2)$ in which σ_2 is pure. There exists a PE σ^p in which $U_{1,\omega}(\sigma_{1,\omega}^p, \sigma_2^p) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2)$ holds for any $\omega \in \Omega$ and σ_2^p is pure and the beliefs of the short-run players are common knowledge on the equilibrium path.*

Proof. See Appendix A.1. □

We show a sketch of the proof. Fix a Nash equilibrium $\sigma = (\sigma_1, \sigma_2)$ in which σ_2 is pure. Given any public history h^t on the equilibrium path, player 1 can infer precisely the player 2's private history h_2^t on the equilibrium path because player 2 uses pure strategy. Full support assumption ensures that any public history h^t is on the equilibrium path. For any public history $h^t = (x^s, y^s)_{s=0}^{t-1}$, let $z_2(h^t)$ be a private history $h_2^t = (\tilde{x}^s, \tilde{y}^s, \tilde{a}_2^s)_{s=0}^{t-1}$ such that h_2^t is on the equilibrium path, and $(\tilde{x}^s, \tilde{y}^s) = (x^s, y^s)$ for $s = 0, 1, \ldots, t-1$. We define public strategy of player 2 as σ_2^p such that $\sigma_2^p(h^t) \equiv \sigma_2(z_2(h^t))$ for any h^t and for any t. The strategies σ_2^p and σ_2 might prescribe different actions off the equilibrium path of σ. However, they prescribe the same action on the equilibrium path of σ. Full support assumption ensures that any private history of player 2 off the equilibrium path is not realized even when player 1 changes his strategies. Hence, for any σ'_1 and $\omega \in \Omega$, $U_{1,\omega}(\sigma'_{1,\omega}, \sigma^p_2) = U_{1,\omega}(\sigma'_{1,\omega}, \sigma_2)$ holds. Thus, σ_1 is a best response to σ^p_2 because (σ_1, σ_2) is a Nash equilibrium. In addition, player 2 is short-run player and we focus on Nash equilibria. Thus, player 2 does not care about a strategy off the equilibrium path. Hence, strategy σ^p_2 is best response to strategy σ_1 because (σ_1, σ_2) is a Nash equilibrium. That is, strategy profile (σ_1, σ^p_2) is also a Nash equilibrium.

Given Nash equilibrium (σ_1, σ^p_2), we define the following public strategy.

$$\sigma^p_{1,\omega}(h^t, x^t) \equiv \int_{\tilde{h}^t_1 \in \mathcal{H}^t_1} \sigma_{1,\omega}(\tilde{h}^t_1, x^t) dg(\tilde{h}^t_1 | \sigma_{1,\omega}, \sigma^p_2, h^t, x^t), \quad \forall \omega \in \Omega,$$

where g is a conditional probability measure on $\mathcal{H}^t_i$ given (σ_1, σ^p_2), h^t and x^t. This means that $U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2) = U_{1,\omega}(\sigma_{1,\omega}, \sigma^p_2)$ for any ω. It implies that σ^p_1 is a best reply to σ^p_2. The public strategy $\sigma^p_1(h^t)$ is a conditional expectation of $\sigma_1(h^t_1)$ given (σ_1, σ_2), h^t and x^t. Hence, σ^p_2 is a best response to σ^p_1 because conditional expectation of $E_{h^t_1}[\sigma_1(h^t_1)|h^t]$ and $\sigma^p_1(h^t)$ are the same for any h^t. Hence, (σ^p_1, σ^p_2) is a Nash equilibrium.

Based on the above discussions, it also holds that $U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2)(= U_{1,\omega}(\sigma_{1,\omega}, \sigma^p_2)) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2)$ for any ω.

Player 1 can keep track of the posterior belief of player 2 on the equilibrium path because player 1 can infer precisely the player 2's private history on the equilibrium path.

3.4 Public equilibrium

It has been shown that, for any Nash equilibrium σ in which σ_2 is pure, the Nash equilibrium payoff can be achieved by a public strategy in which the beliefs of short-run players on the equilibrium path are common knowledge. Based on the results of the previous section, we focus on a public strategy without loss of generality in the sense of Proposition 3.1. In this section, we analyze public equilibria instead of Nash equilibria.

Let us denote by a function $b_1 : \Omega \times [0,1] \to \Delta(A_1)$ player 1's behavior strategy, and we let $b_2 : [0,1] \to A_2$ denote player 2's behavior strategy. For any $\zeta \in \Delta(\Omega)$, function b_1 and b_2, we define a function $T : \Delta(\Omega) \times \Delta(A_1)^{\{\Omega \times [0,1]\}} \times A_2^{[0,1]} \times [0,1] \times Y \to \Delta(\Omega)$ as follows.

$$T(\zeta, b_1, b_2, x, y)(\omega) = \frac{\rho(y|b_1(\omega, x), b_2(x))\zeta(\omega)}{\sum_{\omega' \in \Omega} \rho(y|b_1(\omega', x), b_2(x))\zeta(\omega')}, \quad \forall x \in [0,1], \forall y \in Y, \forall \omega \in \Omega.$$

That is, T is a function that prescribes the Bayesian updating of ζ when public signals x and y are realized given the behavior strategy profile (b_1, b_2). We define $\mathbb{R}$ as the set of real numbers.

We define the set of pairs of the equilibrium payoffs and common priors in which the strategy of player 2 is pure.

$$V = \left\{ (v, \mu) \in \mathbb{R}^\Omega \times \Delta(\Omega) \;\middle|\; \begin{array}{l} \text{Given common prior } \mu \in \Delta(\Omega), \text{there exists a PE } \sigma^p \\ \text{in which } U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2) = v(\omega) \quad \text{for any } \omega \in \Omega, \\ \text{and } \sigma^p_2 \text{ is pure.} \end{array} \right\}$$

For some games, the set V might be empty. One of sufficient conditions that the set V is nonempty is that the set of types of player 1 is singleton and the stage-game has a pure strategy Nash equilibrium.

Let us denote by $W \subset \mathbb{R}^\Omega \times \Delta(\Omega)$ a subset of the set of pairs of equilibrium payoffs and common priors.

Definition 3.2 Fix a subset $W \subset \mathbb{R}^\Omega \times \Delta(\Omega)$. A pair of a function $R : \Omega \to \mathbb{R}$ and a probability distribution $\zeta \in \Delta(\Omega)$ is *decomposable* over W if there exist a behavior strategy $b_1 : \Omega \times [0,1] \to \Delta(A_1)$, a behavior strategy $b_2 : [0,1] \to \Delta(A_2)$

and $\gamma : \Omega \times [0,1] \times Y \to \mathbb{R}$, such that

$$R(\omega) = E_X \left[(1-\delta) u_{1,\omega}(b_1(\omega, X), b_2(X)) + \delta \sum_{y \in Y} \rho(y | b_1(\omega, X), b_2(X)) \gamma(\omega, X, y) \right], \quad \forall \omega \in \Omega, \tag{3.1}$$

$$(1-\delta) u_{1,\omega}(b_1(\omega, x), b_2(x)) + \delta \sum_{y \in Y} \rho(y | b_1(\omega, x), b_2(x)) \gamma(\omega, x, y)$$
$$\geq (1-\delta) u_{1,\omega}(a_1, b_2(x)) + \delta \sum_{y \in Y} \rho(y | a_1, b_2(x)) \gamma(\omega, x, y), \quad \forall a_1 \in A_1, \forall x \in [0,1], \forall \omega \in \Omega, \tag{3.2}$$

$$b_2(x) \in \arg\max_{a_2' \in A_2} \sum_{\omega \in \Omega} u_2(b_1(\omega, x), a_2') \zeta(\omega), \quad \forall x \in [0,1], \tag{3.3}$$

$$(\gamma(\cdot, x, y), T(\zeta, b_1, b_2, x, y)) \in W, \quad \forall x \in [0,1], \forall y \in Y. \tag{3.4}$$

The first condition describes player i's payoff depending on his type when the continuation payoff from next period is given by $\gamma(\omega, x, y)$. The second and third conditions require that player i (weakly) prefers b_i.

The following operator is analogous to the B operator in Cole and Kocherlakota (2001).

Definition 3.3 Let us denote the set of (R, ζ) which is decomposable over W by $B(W)$. A set W is self-generating if $W \subseteq B(W)$ holds.

The following theorem helps us examine whether a set W is a subset of the set V or not.

Theorem 3.1 *Suppose that Assumption 3.1 is satisfied. If a set W is bounded and self-generating, then $W \subseteq V$ holds.*

Proof. By the definition of $B(W)$, any $(R, \zeta) \in B(W)$ is decomposable over W. Hence, for any $(R, \zeta) \in B(W)$, there exist functions b_1, b_2 and γ that satisfy Definition 3.2. We consider functions $\tilde{b}_1 : B(W) \to \Delta(A_1)^{\Omega \times [0,1]}$, $\tilde{b}_2 : B(W) \to \Delta(A_2)^{[0,1]}$ and $\tilde{\gamma} : B(W) \to \mathbb{R}^{\Omega \times [0,1] \times Y}$ such that, for any $(R, \zeta) \in B(W)$, the functions prescribe b_1, b_2 and γ respectively that satisfy Definition 3.2 .

Fix any $(R,\zeta) \in B(W)$. For any public history $h \in \mathcal{H}$, we define a pair of a function $\tilde{R}^h : \Omega \to \mathbb{R}$ and a probability distribution $\tilde{\zeta}^h \in \Delta(\Omega)$ to construct equilibrium strategy. First, we define $\tilde{R}^{h^0} \equiv R$, $\tilde{\zeta}^{h^0} \equiv \zeta$. By the definition of $(R,\zeta) \in B(W)$, we have $(\tilde{R}^{h^0}, \tilde{\zeta}^{h^0}) \in B(W)$.

Next, we fix $t \geq 0$ and consider period $t+1$. Fix a public history in period t, h^t. Suppose that $(\tilde{R}^{h^t}, \tilde{\zeta}^{h^t}) \in B(W)$. For any x^t and y^t, we define $\tilde{R}^{h^t \circ (x^t,y^t)}$ and $\tilde{\zeta}^{h^t \circ (x^t,y^t)}$ when public history $h^t \circ (x^t, y^t)$ is realized in period $t+1$ as follows.

$$\begin{aligned}
\tilde{R}^{h^t \circ (x^t,y^t)} \equiv& \tilde{\gamma}(\tilde{R}^{h^t}, \tilde{\zeta}^{h^t})(\cdot, x^t, y^t), \\
\tilde{\zeta}^{h^t \circ (x^t,y^t)} \equiv& T(\tilde{\zeta}^{h^t}, \tilde{b}_1(\tilde{R}^{h^t}, \tilde{\zeta}^{h^t}), \tilde{b}_2(\tilde{R}^{h^t}, \tilde{\zeta}^{h^t}), x^t, y^t)).
\end{aligned}$$

We have $(\tilde{R}^{h^t \circ (x^t,y^t)}, \tilde{\zeta}^{h^t \circ (x^t,y^t)}) \in B(W)$. Finally, we define public strategy: $\sigma^p_{1,\omega}(h^t, x) \equiv \tilde{b}_1(\tilde{R}^{h^t}, \tilde{\zeta}^{h^t})(\omega, x)$ and $\sigma^p_2(h^t, x) \equiv \tilde{b}_2(\tilde{R}^{h^t}, \tilde{\zeta}^{h^t})(x)$.

By the definition of $\tilde{R}$, for any $h^t \in \mathcal{H}$ and for any $\omega \in \Omega$, it holds that

$$\begin{aligned}
&E_X\left[(1-\delta)u_{1,\omega}(\sigma^p_{1,\omega}(h^t, X), \sigma^p_2(h^t, X))\right] \\
=&\tilde{R}^{h^t}(\omega) - \delta E_X\left[\sum_{y \in Y} \rho(y|\sigma^p_{1,\omega}(h^t, X), \sigma^p_2(h^t, X))\tilde{R}^{h^t \circ (X,y)}(\omega)\right].
\end{aligned}$$

Let us define β^t as a probability measure on $\mathcal{H}^t$ given σ^p. The set W is bounded. Hence, for any $\omega \in \Omega$, $\sum_{t=0}^{\infty} \delta^t \int_{h^t \in \mathcal{H}^t} \tilde{R}^{h^t}(\omega) d\beta^t(h^t|\sigma^p_{1,\omega}, \sigma^p_2)$ is bounded.

Thus, we obtain $U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2)$ as follows.

$$
\begin{aligned}
&U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2) \\
=&\sum_{t=0}^{\infty} \delta^t \int_{h^t \in \mathcal{H}^t} E_X \left[(1-\delta) u_{1,\omega}(\sigma^p_{1,\omega}(h^t, X), \sigma^p_2(h^t, X))\right] d\beta^t(h^t | \sigma^p_{1,\omega}, \sigma^p_2) \\
=&\sum_{t=0}^{\infty} \delta^t \int_{h^t \in \mathcal{H}^t} \left\{ \tilde{R}^{h^t}(\omega) - \delta E_X \left[\sum_{y \in Y} \rho(y | \sigma^p_{1,\omega}(h^t, X), \sigma^p_2(h^t, X)) \tilde{R}^{h^t \circ X \circ y}(\omega) \right] \right\} d\beta^t(h^t | \sigma^p_{1,\omega}, \sigma^p_2) \\
=&\sum_{t=0}^{\infty} \left\{ \int_{h^t \in \mathcal{H}^t} \delta^t \tilde{R}^{h^t}(\omega) d\beta^t(h^t | \sigma^p_{1,\omega}, \sigma^p_2) - \int_{h^{t+1} \in \mathcal{H}^{t+1}} \delta^{t+1} \tilde{R}^{h^{t+1}}(\omega) d\beta^t(h^{t+1} | \sigma^p_{1,\omega}, \sigma^p_2) \right\} \\
=&\sum_{t=0}^{\infty} \int_{h^t \in \mathcal{H}^t} \delta^t \tilde{R}^{h^t}(\omega) d\beta^t(h^t | \sigma^p_{1,\omega}, \sigma^p_2) - \sum_{t=0}^{\infty} \int_{h^{t+1} \in \mathcal{H}^{t+1}} \delta^{t+1} \tilde{R}^{h^{t+1}}(\omega) d\beta^t(h^{t+1} | \sigma^p_{1,\omega}, \sigma^p_2) \\
=&\int_{h^0 \in \mathcal{H}^0} \tilde{R}^{h^0}(\omega) d\beta^0(h^0 | \sigma^p_{1,\omega}, \sigma^p_2) \\
&+ \sum_{t=1}^{\infty} \int_{h^t \in \mathcal{H}^t} \delta^t \tilde{R}^{h^t}(\omega) d\beta^t(h^t | \sigma^p_{1,\omega}, \sigma^p_2) - \sum_{t=0}^{\infty} \int_{h^{t+1} \in \mathcal{H}^{t+1}} \delta^{t+1} \tilde{R}^{h^{t+1}}(\omega) d\beta^t(h^{t+1} | \sigma^p_{1,\omega}, \sigma^p_2) \\
=&\tilde{R}^{h^0}(\omega) = R(\omega).
\end{aligned}
$$

It is proved that, for any $\omega \in \Omega$, $U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2) = R(\omega)$.

In the same way, we find that the continuation payoff of player 1 of type ω for public history h^t is equal to $\tilde{R}^{h^t}(\omega)$. Thus, for any ω and $h^t \in \mathcal{H}$, any one shot deviation at h^t from $\sigma^p_1(\omega, h^t)$ is not profitable for player 1 of type ω by the definition of $\tilde{R}$. Thus, σ^p_1 is the best response to σ^p_2.

Finally, we consider the best response of player 2. By the construction of σ^p, for any $h^t \in \mathcal{H}$, the beliefs $\eta_2(\omega | h^t, \sigma^p)$ is equal to $\tilde{\zeta}(h^t)$. Hence, $\sigma^p_2(h^t)$ is the best response to σ^p_1 given $\eta_2(\omega | h^t, \sigma^p)$ by definition of $\tilde{b}_2$. It has been proved that σ^p is a Nash equilibrium. □

The next theorem shows that the set V is one of the fixed points of operator B.

Proposition 3.2 *Suppose that Assumption 3.1 is satisfied. The set V is bounded and self-generating.*

Proof. The set V is bounded because Ω and A are finite sets. We show that $V \subseteq B(V)$. Fix any $(v, \mu) \in V$ and a PE strategy profile σ^p such that $U_{1,\omega}(\sigma_{1,\omega}, \sigma_2) = v(\omega)$ for any $\omega \in \Omega$ and σ_2^p is pure. Let $U_{1,\omega}(\sigma_{1,\omega}^p, \sigma_2^p | x, y)$ be the expected continuation payoff of player 1 of type ω at $h^1 = (x, y)$. We consider the following b_1, b_2 and γ.

$$\begin{aligned} b_1(\omega, x) =& \sigma_{1,\omega}^p(h^0, x), \\ b_2(x) =& \sigma_2^p(h^0, x), \\ \gamma(\omega, x, y) =& U_{1,\omega}(\sigma_{1,\omega}^p, \sigma_2^p | x, y). \end{aligned}$$

Condition (3.1)-(3.3) in Definition 3.2 are satisfied because σ is a Nash equilibrium. For any $(x, y) \in \mathcal{H}^1$, a pair of continuation payoffs $\gamma(\cdot, x, y)$ and $T(\mu, b_1, b_2, x, y)$ are in V because σ^p is a PE. That is, Condition (3.4) is satisfied. Therefore, we obtain $V \subseteq B(V)$. □

The following proposition shows a monotonicity of B in the sense of set inclusion. It is shown that the set V is the greatest fixed point of operator B.

Proposition 3.3 *Suppose that Assumption 3.1 is satisfied. For any W_1 and W_2, if $W_1 \subseteq W_2$, then $B(W_1) \subseteq B(W_2)$. The set V is the largest set among the bounded and self-generating sets.*

Proof. Fix $w \in B(W_1)$. By definition, w is decomposable over $W_1 (\subseteq W_2)$. Hence, $w \in B(W_2)$. Let $(W_\lambda)_{\lambda \in \Lambda}$ be a family of the set that is bounded and self-generating. For any $\lambda \in \Lambda$, $W_\lambda \subseteq V$ by Theorem 3.1. By Proposition 3.2, the set V is bounded and self-generating. □

The following theorem shows a technique to derive a tight bound of the set V. In addition, the theorem ensures that we can obtain the set V by iterating the technique.

Theorem 3.2 *Suppose that Assumption 3.1 is satisfied. Let a set W be a bounded superset of the set V. Let us define $W_0 \equiv W$, and for $n = 1, 2 \ldots$, $W_n \equiv B(W_{n-1}) \cap W_{n-1}$. Then $\{W_n\}$ is a decreasing sequence and $\lim_{n \to \infty} W_n = V$.*

Proof. Fix any set W such that $W \supseteq V$. By definition, $\{W_n\}$ is a decreasing sequence. Thus, $\lim_{n\to\infty} W_n$ is $\cap_{n\in\mathbb{N}} W_n$. For any n, $W_{n+1} \subseteq B(W_n)$. Hence, we obtain

$$\lim_{n\to\infty} W_n = \lim_{n\to\infty} W_{n+1} \subseteq \lim_{n\to\infty} B(W_n) = B\left(\lim_{n\to\infty} W_n\right)$$

Thus, $\lim_{n\to\infty} W_n$ is self-generating. In addition, it holds that

$$W = W_0 \supseteq W_1 \supseteq W_2 \cdots$$

Thus, $\lim_{n\to\infty} W_n$ is bounded. Hence, we obtain $\lim_{n\to\infty} W_n \subseteq V$ by Theorem 3.1.

Next, we show that $\lim_{n\to\infty} W_n \supseteq V$. By Proposition 3.3, for any set W', if the set W' is a superset of V, then it holds that

$$B(W') \supseteq B(V) = V$$

Thus, for any W', if W' is a superset of V, then it holds that $B(W') \cap W' \supseteq V$.

Hence, if a set W_n is a superset of the set V, then the set W_{n+1} is also a superset of the set V. Hence, we obtain that $W_n \supseteq W$ for any n because the set W_0 is a superset of the set V. Therefore, $\lim_{n\to\infty} W_n \supseteq V$ has been shown. □

Let us consider a set $\mathbb{R}^\Omega \times \Delta(\Omega)$ and Euclidean distance over $\mathbb{R}^\Omega \times \Delta(\Omega)$.. Let us denote by $\mathrm{Cl}(W)$ the closure of the set W. The following lemma shows a property of $B(W)$.

Lemma 3.1 *Suppose that Assumption 3.1 is satisfied. If a set W is compact, then it holds that* $\mathrm{Cl}(B(W)) = B(W)$.

Proof. Fix any convergent sequence $(R^s, \zeta^s)_{s=1}^\infty$ such that, for any $s \in \mathbb{N}$, $(R^s, \zeta^s) \in W$. We let R denote $\lim_{s\to\infty} R^s$ and let ζ denote $\lim_{s\to\infty} \zeta^s$. For any $s \in \mathbb{N}$, there exist b_1^s, b_2^s and γ^s such that (R^s, ζ^s) is decomposable over the set W with b_1^s, b_2^s and γ^s. For any s and x and w, $\gamma(\omega, x, w)$ is a continuation payoff for type ω of player 1. Both the set of action profiles a and the set of types are finite. Hence, $\gamma(\omega, x, w)$ is bounded from below by the minimum expected stage-game payoff $\underline{u} \equiv \min_{\omega\in\Omega} \min_{a\in A} u_1(\omega, a)$ and bounded from above by the maximum expected payoff $\overline{u} \equiv \max_{\omega\in\Omega} \max_{a\in A} u_1(\omega, a)$. Thus, the sequence $(b_1^s, b_2^s, \gamma^s)_{s=1}^{\infty}$ is a sequence in the compact set $\left(\Delta(A_1)^{\Omega\times[0,1]}\right) \times \left(A_2^{[0,1]}\right) \times \left([\underline{u}, \overline{u}]^{\Omega\times[0,1]\times Y}\right)$. Therefore, we can choose the convergent subsequence $(b_1^{s_\ell}, b_2^{s_\ell}, \gamma^{s_\ell})_{\ell=1}^{\infty}$. We let b_1 denote $\lim_{\ell\to\infty} b_1^{s_\ell}$, let b_2 denote $\lim_{\ell\to\infty} b_2^{s_\ell}$ and let by γ denote $\lim_{\ell\to\infty} \gamma^{s_\ell}$. By construction and the upper hemicontinuity of the best response, (R, ζ) is decomposable over W with b_1, b_2 and γ. Thus, we obtain $(R, \zeta) \in W$. □

Finally, we show a compactness of the set V.

Proposition 3.4 *Suppose that Assumption 3.1 is satisfied. The set V is compact.*

Proof. By Proposition 3.3, we obtain $V = B(V) \subseteq B(\mathrm{Cl}(V))$. By Lemma 3.1, it holds that $\mathrm{Cl}(V) \subseteq \mathrm{Cl}(B(\mathrm{Cl}(V))) = B(\mathrm{Cl}(V))$ because $\mathrm{Cl}(V)$ is compact. Therefore, $\mathrm{Cl}(V)$ is bounded and self-generating. Hence, we obtain $\mathrm{Cl}(V) \subseteq V$. It holds that $\mathrm{Cl}(V) \supseteq V$ by the definition of closure. Thus, we obtain $\mathrm{Cl}(V) = V$. Hence, it is proved that the set V is compact. □

3.5 Example : Participation game

In this section, we analyze a *Participation game*. The participation game is defined as follows. There exist a firm (Player 1) and a consumer (Player 2). Player 1 decides whether to exert high effort to provide a service for player 2 or not. Player 2 decides whether to use the service or not. Player 1 and player 2 decide their actions simultaneously. If player 2 does not use the service (D), both players obtain payoff 0. If player 1 exerts high effort (H) and player 2 uses the service (U), then both players obtain expected payoff 1. If player 1 does not exert high effort (L) and player 2 uses the service, then player 1 obtains expected payoff $1 + c$ and player 2 obtains expected payoff -1. The expected stage-game payoff matrix is summarized as follows.

Participation game

		Player 2	
		U	D
Player 1	H	$1, 1$	$0, 0$
	L	$1 + c, -1$	$0, 0$

Players observe a public signal. The set of public signals is given by $Y \equiv \{y_H, y_L\}$. The distribution of public signal y is as follows.

$$\rho(y_H|H, U) = 1 - \rho(y_H|L, U) = p,$$
$$\rho(y_H|L, D) = \rho(y_H|H, D) = 1 - q.$$

We assume that $p > \frac{1}{2}$. That is, when player 1 chooses action a_1 and player 2 chooses action U, a public signal y_{a_1}, which corresponds to the action chosen by player 1, is realized with high probability $p(> \frac{1}{2})$ and signal $y_{a_1'}$ is realized with low probability $1 - p(< \frac{1}{2})$. However, when player 2 chooses action D, the distribution of public signals does not depend on player 1's action.

We consider an infinitely repeated participation game. Player 1 is a firm with a fixed discount factor $\delta \in [0, 1)$ and player 2 is a sequence of short-run consumers. At the beginning of the infinitely repeated game, Nature selects player 1's type from $\{\omega^*, \omega_0\}$. With probability $\mu > 0$, player 1 is a normal type (ω^*) and he chooses an action from $\{H, L\}$ in each period. With probability $1 - \mu > 0$, player 1 is a bad type (ω_0), and then the repetition of action L is a dominant-action of the infinitely repeated game. We assume that public randomization device is available.

We cannot use the results of Cole and Kocherlakota (2001) in the above participation game. A strategy profile is *Markov-private* if each player's strategy depends on his past private information only through his current private state. Given a game, beliefs are said to be *Markov* if, for any Markov-private strategy profile, each player's belief about the other players' states depend on his private information only through his current state, both on and off the equilibrium path. Cole and Kocherlakota (2001) present a technique to derive a set of the equilibrium payoff vectors and common priors over the states in a game if beliefs are Markov in the game.

Beliefs are not Markov in the above participation game. Consider the following strategies. Normal type chooses action H at any history, and bad type chooses action L at any history. Player 2 chooses action D at any history. These strategies are Markov-private. If player 2 plays action D and y_H is realized in period 0, then posterior belief over player 1's types in period 1 is equal to common prior over player 1's types. However, if player 2 deviates to play action U and y_H is realized in period 0, then posterior belief over player 1's types in period 1 defers from common prior over player 1's types. That is, it does not hold that player 2's beliefs about player 1's types depends on his private information only through his current state. Thus, we cannot use the results of Cole and Kocherlakota (2001).

We study the set of pairs of the equilibrium payoff vectors and common priors in which player 2's strategy is pure by using our results from Section 3.4. First, we provide a candidate of the set of pairs of the equilibrium payoff vectors and common priors. Second, we show that the candidate is bounded and self-generating. Thus, it is proved by Theorem 3.1 that the candidate is a subset of the set of pairs of equilibrium payoffs and common priors.

The bad type is a commitment type. Therefore, we do not show the payoff for the bad type. We focus on payoffs of the normal type, and we consider a set of pairs of equilibrium payoffs for normal type and a common prior probability with which player 1 is normal type, $W \subseteq \mathbb{R} \times [0, 1]$.

3.5.1 A candidate of the set of pairs of equilibrium payoffs and common priors

First, to define a candidate of the set of pairs of equilibrium payoffs and common priors W, we define ψ, ξ and $\overline{v}$ as follows.

$$\begin{aligned}\psi \equiv& \frac{1+\sqrt{1-4\delta^2 p(1-p)}}{2\delta p} > 1\\ \xi \equiv& \frac{1-\sqrt{1-4\delta^2 p(1-p)}}{2\delta p} < 1\\ \overline{v} \equiv& 1-\frac{1-p}{2p-1}c\end{aligned}$$

Second, for any nonnegative integers N, we define α_N, β_N as a solution to the following simultaneous linear equations.

$$1+\alpha_N\psi^N+\beta_N\xi^N = (1-\delta)+\delta p\overline{v}+\delta(1-p)\left[1+\alpha_N\psi^{N-1}+\beta_N\xi^{N-1}\right] \tag{3.5}$$

$$1+\alpha_N+\beta_N = (1-\delta)+\delta p\left(1+\alpha_N\psi+\beta_N\xi\right) \tag{3.6}$$

Let us define $v_N : \mathbb{Z} \to \mathbb{R}$ as follows.

$$v_N(s) = \begin{cases} \overline{v} & \text{if } s > N, \\ 1+\alpha_N\psi^s+\beta_N\xi^s, & \text{if } 0 \le s \le N, \\ 0 & \text{if } s < 0. \end{cases}$$

We define N^*.

$$N^* \equiv \min\left\{ N \in \mathbb{N}\cup\{0\} \;\middle|\; \begin{array}{l} v_N(s) \text{ is non-decreasing function of } s, \text{ and} \\ \overline{v}-V_N(N) \le \frac{1-\delta}{\delta}\frac{c}{2p-1}. \end{array} \right\}$$

In some games, the above set is empty. Then, there exists no nonnegative integer N^*.

Finally, we define I_s, E and W as follows, where $\mathbb{Z}$ denotes the set of integers.

$$I_s \equiv \left[\frac{p^s}{p^s+(1-p)^s}, \frac{p^{s+1}}{p^{s+1}+(1-p)^{s+1}}\right),$$

$E \equiv \{(v,\mu)|$ Given μ, there exists pure strategy Nash equilibrium in which ω^* obtains payoff v.$\}$,

$W \equiv \{(v,\mu)| v \in [0, v_{N^*}(s)]$ and $\mu \in I_s$ for some $s \in \mathbb{Z}\}$.

The set is W is well defined only if there exists N^*.

Given any $\mu \in I_s$, when normal type chooses H and y_H (resp. y_L) is realized given μ, then Bayesian updated μ is in I_{s+1} (resp. I_{s-1}). This statement is summarized as follows.

Fact 3.1 *If $\mu \in I_s$, then it holds that*

$$\frac{\mu p}{\mu p + (1-\mu)(1-p)} \in I_{s+1}, \quad \text{and} \quad \frac{\mu(1-p)}{\mu(1-p) + (1-\mu)p} \in I_{s-1}.$$

The expression $\frac{\mu p}{\mu p+(1-\mu)(1-p)}$ (resp. $\frac{\mu(1-p)}{\mu(1-p)+(1-\mu)p}$) is Bayesian updated μ when the normal type chooses H and y_H (resp. y_L) is realized.

If $\mu \in I_s$ for some negative integer s, then the short-run player's best response is action D even if the normal type chooses action H. If $\mu \in I_s$ for some nonnegative integer s, then the short-run player's best response is action U if and only if the normal type chooses action H.

3.5.2 A subset of the set of pairs of equilibrium payoffs and common priors

In general, we need a complicated proof to show that any element of the set W is a pair of the equilibrium payoff and common prior. However, our result enables us to show it easily. In this subsection, we show that $W \subseteq E$ by Theorem 3.1.

Proposition 3.5 *Fix any discount factor $\delta \in [0,1)$. If there exists N^*, then $W \subseteq E$.*

If there exists no N^*, then a set W is not well defined. To prove Proposition 3.5, we show the following lemmas 3.2–3.3. We do not use public randomization to prove the following lemmas. Hence, we omit public randomization x from the definitions of T, b, γ for simplicity.

Using substitution of $v_{N^*}(s)$, we have the following lemma.

Lemma 3.2 *If there exists N^*, then it holds that*

$$v_{N^*}(s) = \begin{cases} \overline{v} & \text{if } s > N^*, \\ (1-\delta) + \delta p v_{N^*}(s+1) + \delta(1-p)v_{N^*}(s-1) & \text{if } 0 \leq s \leq N^*, \\ 0 & \text{if } s < 0. \end{cases}$$

Proof. If $s > N^*$ holds, then we obtain $v_{N^*}(s) = \overline{v}$ by the definition of $v_{N^*}(s)$. If $s = N^*$ holds, then we obtain $v_{N^*}(s) = (1-\delta) + \delta p v_{N^*}(s+1) + \delta(1-p)v_{N^*}(s-1)$ by equation (3.5). In the same way, If $s < 0$ holds, then we obtain $v_{N^*}(s) = 0$ by the definition of $v_{N^*}(s)$. If $s = 0$ holds, then we obtain $v_{N^*}(s) = (1-\delta) + \delta p v_{N^*}(s+1) + \delta(1-p)v_{N^*}(s-1)$ by equation (3.6). Finally, we consider $s = 1, 2, \ldots N^* - 1$.

$$\begin{aligned} &(1-\delta) + \delta p v_{N^*}(s+1) + \delta(1-p)v_{N^*}(s-1) \\ =&1 + \delta p(\alpha_{N^*}\psi^{s+1} + \beta_{N^*}\xi^{s+1}) + \delta(1-p)(\alpha_{N^*}\psi^{s-1} + \beta_{N^*}\xi^{s-1}) \\ =&1 + \alpha_{N^*}\psi^s + \beta_{N^*}\xi^s. \end{aligned}$$

The last equality follows from the fact that ψ and ξ are theolution of the following quadratic formula.

$$\delta p x^2 + \delta(1-p) = x.$$

□

From the above lemma, we obtain the following lemma.

Lemma 3.3 *If there exists N^*, then, for any $s \in \{0, 1, 2, \ldots, N^* - 1\}$, it holds that*

$$v_{N^*}(s) - v_{N^*}(s-1) > v_{N^*}(s+1) - v_{N^*}(s) > \frac{1-\delta}{\delta}\frac{c}{2p-1}.$$

Proof. By lemma 3.2, for any $s \in \{0, 1, 2, \ldots, N^*\}$, it holds that

$$v_{N^*}(s) = (1-\delta) + \delta p v_{N^*}(s+1) + \delta(1-p)v_{N^*}(s-1).$$

From the above equation, we have

$$\begin{aligned} v_{N^*}(s+1) - v_{N^*}(s) =& \frac{1-p}{p}(v_{N^*}(s) - v_{N^*}(s-1)) - \frac{1-\delta}{\delta p}(1 - v_{N^*}(s)) \\ <& v_{N^*}(s) - v_{N^*}(s-1). \end{aligned}$$

The last inequality follows from $p > \frac{1}{2}$ and $v_{N^*}(s) \leq \overline{v} < 1$. Hence, we obtain $v_{N^*}(s) - v_{N^*}(s-1) > v_{N^*}(s+1) - v_{N^*}(s)$.

If $\overline{v} - v_{N^*}(N^* - 1) \leq \frac{1-\delta}{\delta}\frac{c}{2p-1}$ holds, then $\overline{v} - v_{N^*-1}(N^* - 1) \leq \frac{1-\delta}{\delta}\frac{c}{2p-1}$ and $v_{N^*-1}(s)$ is a nondecreasing function of s. This is a contradiction to the definition of N^*. Therefore, by the definition of N^*, it holds that

$$\overline{v} - v_{N^*}(N^* - 1) = v_{N^*}(N^*) - v_{N^*}(N^* - 1) > \frac{1-\delta}{\delta}\frac{c}{2p-1}.$$

□

Using the previous lemma, we obtain the following lemma.

Lemma 3.4 *If there exists N^*, then, for any $s \in \{0, 1, \dots\}$, it holds that*

$$\begin{aligned} &\min\left\{v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1}, v_{N^*}(s-1)\right\} \\ =& \begin{cases} v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1} & \text{if } s > N^*, \\ v_{N^*}(s-1) & \text{if } 0 \leq s \leq N^*. \end{cases} \end{aligned}$$

Proof. The function $v_{N^*}(s)$ is non-decreasing function of s. Hence, by Lemma 3.3, it holds that

$$v_{N^*}(s+1) - v_{N^*}(s-1) \geq v_{N^*}(s) - v_{N^*}(s-1) > \frac{1-\delta}{\delta}\frac{1}{2p-1}, \quad \forall s \in \{0, 1, 2, \dots, N^*\}.$$

Hence, for any $s \in \{0, 1, 2, \dots, N^*\}$, it holds that

$$\min\left\{v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1}, v_{N^*}(s-1)\right\} = v_{N^*}(s-1)$$

By the definition of N^*, it holds that, for any $s \in \{N^*, N^* + 1, \dots\}$

$$v_{N^*}(s+1) - v_{N^*}(s) = \overline{v} - v_{N^*}(s) \leq \frac{1-\delta}{\delta}\frac{c}{2p-1}.$$

Hence, if there exists N^*, then, for any $s \in \{0, 1, \dots\}$, it holds that

$$\begin{aligned}
&\min\left\{v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1}, v_{N^*}(s-1)\right\} \\
&= \begin{cases} v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1} & \text{if } s > N^*, \\ v_{N^*}(s-1) & \text{if } 0 \leq s \leq N^*. \end{cases}
\end{aligned}$$

□

Player 1 prefers action H when $\gamma(y_H) - \gamma(y_L) \geq \frac{1-\delta}{\delta}\frac{c}{2p-1}$ holds. If it holds that

$$\begin{aligned}
\gamma(y_H) =& v_{N^*}(s+1), \quad \text{and} \\
\gamma(y_L) =& \min\left\{v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1}, v_{N^*}(s-1)\right\},
\end{aligned}$$

then, it holds that $\gamma(y_H) - \gamma(y_L) \geq \frac{1-\delta}{\delta}\frac{c}{2p-1}$. Hence, from Lemma 3.4, we have the following fact.

Fact 3.2 *If there exists N^*, then, for any nonnegative integer $s \in \{0, 1, 2, \dots\}$ and for any $\mu \in I_s$, it holds that if*

$$\begin{aligned}
&b_1(\omega^*) = H, b_1(\omega_0) = L, b_2 = U, \\
&\gamma(y_H) = v_{N^*}(s+1), \gamma(y_L) = \min\left\{v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1}, v_{N^*}(s-1)\right\},
\end{aligned}$$

then,

$$\begin{aligned}
v_{N^*}(s) =& (1-\delta) + \delta p \gamma(y_H) + \delta(1-p)\gamma(y_B) \\
&\geq (1-\delta)(1+c) + \delta(1-q)\gamma(y_H) + \delta(1-q)\gamma(y_B), \\
&b_2 \in \arg\max_{b_2 \in A_2} \mu u_2(b_1(\omega^*), b_2) + (1-\mu)u_2(b_1(\omega_0), b_2), \\
&(\gamma(y), T(\mu, y)) \in W, \quad \forall y \in Y,
\end{aligned}$$

where $T(\mu, y)$ is Bayesian updated μ when y is realized.

If γ does not change irrespective of the public signals, then player 1 prefers the dominant-action of the stage-game. Hence, we have the following fact.

Fact 3.3 *If there exists N^*, then, for any integer $s \in \mathbb{Z}$ and for any $\mu \in I_s$, it holds that if*

$$\begin{aligned} b_1(\omega^*) =& b_1(\omega_0) = L, b_2 = D, \quad \textit{and} \\ \gamma(y_H) =& \gamma(y_L) = 0 \end{aligned}$$

then,

$$\begin{aligned} 0 =& (1-\delta)0 + \delta q \gamma(y_H) + \delta(1-q)\gamma(y_B) \\ \geq& (1-\delta)0 + \delta q \gamma(y_H) + \delta(1-q)\gamma(y_B), \\ b_2 \in& \arg\max_{b_2 \in A_2} \mu u_2(b_1(\omega^*), b_2) + (1-\mu) u_2(b_1(\omega_0), b_2), \\ (\gamma(y)&, T(\mu, y)) \in W, \quad \forall y \in Y, \end{aligned}$$

where $T(\mu, y)$ is Bayesian updating of μ when y is realized.

Finally, we prove Proposition 3.5.

Proof of Proposition 3.5. For any nonnegative integer $s \in \mathbb{Z}$ and for any $\mu \in I_s$, $(v_{N^*}(s), \mu)$ is decomposable over the set W by b_1, b_2 and γ described in Fact 3.2. In addition, we obtain that for any integer $s \in \mathbb{Z}$ and for any $\mu \in I_s$, $(0, \mu)$ is decomposable over the set W by b_1, b_2 and γ described in Fact 3.3. Hence, for any nonnegative integer $s \in \mathbb{Z}$, any $(v, \mu) \in [0, v_{N^*}(s)] \times I_s$ is decomposable over the set W

by the following b_1, b_2, γ and the following public randomization.

$$
\begin{aligned}
b_1(\omega^*, x) =& \begin{cases} H, & \text{if } x \leq \frac{v}{v_{N^*}(s)}, \\ L, & \text{if } x > \frac{v}{v_{N^*}(s)}, \end{cases} \\
b_1(\omega_0, x) =& L, \\
b_2(x) =& \begin{cases} U, & \text{if } x \leq \frac{v}{v_{N^*}(s)}, \\ D, & \text{if } x > \frac{v}{v_{N^*}(s)}, \end{cases} \\
\gamma(y_H) =& \begin{cases} v_{N^*}(s+1), & \text{if } x \leq \frac{v}{v_{N^*}(s)}, \\ 0, & \text{if } x > \frac{v}{v_{N^*}(s)}, \end{cases} \\
\gamma(y_L) =& \begin{cases} \min\left\{ v_{N^*}(s+1) - \frac{1-\delta}{\delta}\frac{c}{2p-1}, v_{N^*}(s-1) \right\}, & \text{if } x \leq \frac{v}{v_{N^*}(s)}, \\ 0, & \text{if } x > \frac{v}{v_{N^*}(s)}. \end{cases}
\end{aligned}
$$

Hence, the set W is bounded and self-generating. Proposition 3.5 has been proved by Theorem 3.1. □

3.6 Extension to mixed strategies

We now briefly describe how to extend the results of Section 3.4 to the set of pairs of equilibrium payoffs and common priors in which σ_2 is not pure. The key idea in Section 3.3 and Section 3.4 is that if σ_2 is pure, then player 1 can keep track of the beliefs of short-run players on the equilibrium path. Hence, for any Nash equilibrium, we can analyze any subgame on the equilibrium path as if it is a initial period given some common prior.

Standard assumptions that are often made in reputation models ensure that player 1 can keep track of the beliefs of short-run players on the equilibrium path even if σ_2 is not pure.

3.6.1 Observability

Many studies (e.g., Ely et al. (2008) and Kaya (2009)) assume that the actions chosen by a short-run players are observable to each player. In this subsection, we discuss the way to extend our results to the case of mixed strategies under the following assumption.

Assumption 3.2 Player 2's actions are observable.

If player 2's actions are observable, player 1 can know the action chosen by player 2. Thus, player 1 can keep track of the beliefs of short-run players even if player 2 chooses mixed action.

Let $h_1^t = (x^s, a^s, y^s)_{s=0}^{t-1}$ be a private history of player 1 at the beginning of period $t \geq 1$, and $h^t = (x^s, a_2^s, y^s)_{s=0}^{t-1}$ be a public history at the beginning of period $t \geq 1$. First, we consider the following variant of Proposition 3.1 including mixed strategies under Assumption 3.1 and Assumption 3.2.

Proposition 3.6 *Suppose that Assumption 3.1 and Assumption 3.2 are satisfied. Fix common prior* $\mu \in \Delta(\Omega)$ *and a Nash equilibrium* $\sigma = (\sigma_1, \sigma_2)$. *Then, there exists a PE* σ^p *in which* $U_{1,\omega}(\sigma_{1,\omega}^p, \sigma_2^p) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2)$ *holds for any* $\omega \in \Omega$.

A sketch of the proof is as follows. Fix any Nash equilibrium σ. For any public history h^t on the equilibrium path of σ, we define the public strategy σ^p as follows.

$$\begin{aligned}\sigma_{1,\omega}^p(h^t, x^t) \equiv & \int_{\tilde{h}_1^t \in \mathcal{H}_1^t} \sigma_{1,\omega}(\tilde{h}_1^t, x^t) dg(\tilde{h}_1^t | \sigma_{1,\omega}, \sigma_2, h^t, x^t), \quad \forall \omega \in \Omega, \\ \sigma_2^p(h^t, x) \equiv & \sigma_2(x).\end{aligned}$$

For any public history h^t off the equilibrium path of σ, we define $\sigma_{1,\omega}^p(h^t, x^t) \equiv a_1$ for some $a_1 \in A_1$ and $\sigma_2^p(h^t, x) = A_2$ for some $a_2 \in A_2$.

By the construction of σ_2^p and Assumption 3.1, any public history h^t off the equilibrium path of σ never realizes whatever player 1 plays. Hence, the continuation strategy of player 2 off the equilibrium path does not affect player 1's best response on the equilibrium path of σ. In addition, the payoff for player 2 in any period t is not affected by the continuation strategy off the equilibrium path because player 2 is a short-run player. Therefore, we can use a similar argument in the proof of Proposition 3.1. We obtain Proposition 3.6.

Finally, we discuss the rest of our results in Section 3.4 in mixed strategies under Assumption 3.2. If player 2's actions are observable, then a new public signal (a_2, y) is realized instead of y itself in each period. This new public signal does not satisfy Assumption 3.1. Given some Nash equilibrium, some public histories are off the equilibrium path. However, these public history off the equilibrium path is realized if and only if player 2 deviates in the past. In the same way as the proof in Section 3.3, we ignore player 2's deviation because player 2 is short-run players and we focus on Nash equilibria. Hence, we require Condition 3.4 in Definition 3.2 with respect to (a_2, y) only when b_2 prescribes a_2 with positive probability. Other conditions are the same as Section 3.3 except that b_2 is a function of $[0, 1]$ to $\Delta(A_2)$. Then, we can use the same proofs in Section 3.3 and Section 3.4.

3.6.2 Independence

Some studies analyze reputation models without Assumption 3.2. However, these works analyze repeated games under the following assumption to ensure that player 1 can keep track of the belief of player 2 (e.g., Tadelis (1999) Mailath and Samuelson (2001)). In this section, we discuss our results including mixed strategies in such reputation models.

Assumption 3.3 For any actions a_2 and a_2', it holds that

$$\rho(\cdot|\cdot, a_2) = \rho(\cdot|\cdot, a_2').$$

Assumption 3.3 ensures that short-run players' action a_2 contains no useful information. Hence, player 2's best response is the same between two private histories that induce the same public history. Given any Nash equilibrium σ, consider the following public strategy of player 2.

$$\sigma_2^p(h^t, x^t) \equiv \int_{\tilde{h}_2^t \in \mathcal{H}_2^t} \sigma_2(\tilde{h}_2^t, x^t) dg(\tilde{h}_2^t | \sigma, h^t, x^t).$$

That is, $\sigma_2^p(h^t, x^t)$ is an expectation of $\sigma_2(h_2^t, x^t)$. In each period, player 2 chooses a convex combination of mixed actions among best response given σ. Hence, σ_2^p is a best response to σ_1. By the construction of σ_2^p, it holds that $U_{1,\omega}(\sigma'_{1,\omega}, \sigma_2^p) = U_{1,\omega}(\sigma'_{1,\omega}, \sigma_2)$ for any ω and for any σ'. We define σ_1^p in the same way as in the previous subsection. Then, we have $U_{1,\omega}(\sigma_{1,\omega}^p, \sigma_2^p) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2^p) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2)$. Hence, σ_1^p is a best response to σ_2^p because σ is a Nash equilibrium.

By the construction of σ^p and Assumption 3.3, it holds that $E[u_2(\cdot, \cdot)|\sigma^p, h^t] = E[u_2(\cdot, \cdot)|\sigma, h^t]$ for any a_2 and h^t. Hence, σ_2^p is a best response to σ_1^p. It has been proved that σ^p is a PE. We obtain a variant of Proposition 3.1 including mixed strategies under Assumption 3.1 and 3.3.

3.7 Conclusion

The technique in Abreu et al. (1990) is useful for studying the equilibrium payoff sets in infinitely repeated games. In this chapter, we extend the idea to a range of infinitely repeated games with a long-run player who has a persistent type and a sequence of short-run players. Most of studies on reputation models characterize the bounds of the equilibrium payoff set because it is difficult to analyze the set of the equilibrium payoffs. However, we show the following two results about the set of pairs of equilibrium payoffs and common priors.

First, we find a technique of checking whether or not a set of payoff vectors and common priors is a subset of the set of pairs of equilibrium payoffs and common priors (Theorem 3.1). It simplifies proving that any element in a set is attainable as a Nash equilibrium payoff vector. This technique is the most useful when we have a candidate for the set of pairs of the equilibrium payoffs and common priors. However, sometimes, we have no candidate for the set of pairs of the equilibrium payoffs and common priors. Our second result is useful in such a situation. Theorem 3.2 gives us a smaller superset of the set of pairs of the equilibrium payoffs and common priors from a bounded superset of the set of pairs of the equilibrium payoffs and common priors. We show that by iterating this operation, we can find an arbitrary tight superset of the set of pairs of equilibrium payoffs and common priors.

Appendix

A.1 Proof of Proposition 3.1

In this section, we prove Proposition 3.1. First, we show that, for any pure strategy Nash equilibrium, the actions chosen by short-run players on the equilibrium path are common knowledge. Next, we prove Proposition 3.1.

Fix common prior $\mu \in \Delta(\Omega)$ and Nash equilibrium $\sigma = (\sigma_1, \sigma_2)$ in which σ_2 is pure. First, we define public strategy σ^p and, next we show that $U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2)$ holds.

For any public history $\hat{h}^t = (\hat{x}^s, \hat{y}^s)_{s=0}^{t-1}$ and public signal $\hat{x}^t$, we define a sequence of actions chosen by player 2 on the equilibrium path $(\hat{a}_2^s)_{s=0}^{t-1}$, as follows.

$$\hat{a}_2^s = \begin{cases} \sigma_2(h_2^0, \hat{x}^0) & \text{if } s = 0, \\ \sigma_2((\hat{x}^\tau, \hat{a}_2^\tau, \hat{y}^\tau)_{\tau=0}^{s-1}, \hat{x}^s) & \text{if } 0 < s < t. \end{cases}$$

That is, the sequence of actions chosen by short-run players is generated by a public history and a sunspot so that it is common knowledge on the equilibrium path. It means that each player can keep track of the beliefs of short-run players on the equilibrium path.

Then, we can define σ_2^p as follows.

$$\sigma_2^p(\hat{h}^t, \hat{x}) = \begin{cases} \sigma_2(h_2^0, \hat{x}) & \text{if } t = 0 \\ \sigma_2((\hat{x}^s, \hat{a}_2^s, \hat{y}^s)_{s=0}^{t-1}, \hat{x}) & \text{if } t \in \mathbb{N}. \end{cases}$$

For any strategy profile $\tilde{\sigma}$ and for any $t \in \{0\} \cup \mathbb{N}$, let $f(\omega, x^t, y^t, h_1^t, h_2^t, h^t|\tilde{\sigma})$ be the probability measure on $\Omega \times [0,1] \times Y \times \mathcal{H}_1^t \times \mathcal{H}_2^t \times \mathcal{H}^t$ given $\tilde{\sigma}$. For any $t \in \{0\} \cup \mathbb{N}$, let $\mathrm{supp}_\sigma(\mathcal{H}_i^t)$ be a support of private history h_i^t given the strategy profile σ. By the construction of σ_2^p, it holds that $\sigma_2^p(h_2^t, \cdot) = \sigma_2(h_2^t, \cdot)$ for any $h_2^t \in \mathrm{supp}_\sigma(\mathcal{H}_2^t)$. Therefore, for any player 1's strategy $\tilde{\sigma}_1$ and for any $z \in \Omega \times [0,1] \times Y \times \mathcal{H}_1^t \times \mathcal{H}_2^t \times \mathcal{H}^t$, it holds that

$$f(z|\tilde{\sigma}_1, \sigma_2^p) = f(z|\tilde{\sigma}_1, \sigma_2). \tag{3.7}$$

This means that the behavior of short-run players does not affect the long-run player's payoff because for any long run player's strategy, short run players' strategies σ_2^p and σ_2 generate the same distribution of z.

By the full support assumption, any public history $h \in \mathcal{H}$ is realized with positive probability. Hence, for any t, the conditional probability measure g on $\mathcal{H}_i^t$ given σ, h^t and x^t is well-defined. We define public strategy $\sigma_{1,\omega}^p$ as follows.

$$\sigma_{1,\omega}^p(h^t, x^t) \equiv \int_{\tilde{h}_1^t \in \mathcal{H}_1^t} \sigma_{1,\omega}(\tilde{h}_1^t, x^t) dg(\tilde{h}_1^t|\sigma_{1,\omega}, \sigma_2^p, h^t, x^t), \quad \forall \omega \in \Omega.$$

It implies that $\sigma_{1,\omega}^p(h^t, x^t)$ is the expected value of $\sigma_{1,\omega}(\tilde{h}_i^t, x^t) \in \Delta(A_1)$. For any t, we define the conditional probability measure e on the set of y^t, x^{t+1}, a_1^{t+1} given σ, h^t and x^t.

By the construction of the strategy σ_1^p, for any $t \in \{0\} \cup \mathbb{N}$, $h^t \in \mathcal{H}^t$, $x \in [0,1]$, $a_1 \in A_1$, and $y \in Y$, it holds that,

$$e(y^t, x^{t+1}, a_1^{t+1}|\sigma_{1,\omega}^p, \sigma_2^p, h^t, x^t) = e(y^t, x^{t+1}, a_1^{t+1}|\sigma_{1,\omega}, \sigma_2^p, h^t, x^t)$$

For any t, we define the probability measure ℓ on the set of h^t, x^t, a_1^t induced by σ. We show that, for any $\omega \in \Omega$, and for any $t \in \{0\} \cup \mathbb{N}$, it holds that

$$\ell(h^t, x^t, a_1^t|\sigma_{1,\omega}^p, \sigma_2^p) = \ell(h^t, x^t, a_1^t|\sigma_{1,\omega}, \sigma_2^p), \quad \forall h^t \in \mathcal{H}^t,\ \forall x^t \in [0,1],\ \forall a_1 \in A_1 \tag{3.8}$$

We obviously obtain that (3.8) holds for $t = 0$. If (3.8) holds for $t = n$, then (3.8) holds for $t = n+1$, as follows. For any t, let us define the probability measure $\tilde{\beta}^t$ on the set of h^t and x^t induced by σ. Then, $\tilde{\beta}^t(h^t, x^t|\sigma^p_{1,\omega}, \sigma^p_2) = \tilde{\beta}^t(h^t, x^t|\sigma_{1,\omega}, \sigma^p_2)$. Hence,

$$\begin{aligned}
\ell((h^t \circ (x^t, y^t)), x^{t+1}, a_1^{t+1}|\sigma^p_{1,\omega}, \sigma^p_2) =& \tilde{\beta}^t(h^t, x^t|\sigma^p_{1,\omega}, \sigma^p_2) e(y^t, x^{t+1}, a_1^{t+1}|\sigma^p_{1,\omega}, \sigma^p_2, h^t, x^t) \\
=& \tilde{\beta}^t(h^t, x^t|\sigma_{1,\omega}, \sigma^p_2) e(y^t, x^{t+1}, a_1^{t+1}|\sigma_{1,\omega}, \sigma^p_2, h^t, x^t) \\
=& \ell((h^t \circ (x^t, y^t)), x^{t+1}, a_1^{t+1}|\sigma_{1,\omega}, \sigma^p_2).
\end{aligned}$$

Thus, it is shown that (3.8) holds for any $t \in \{0\} \cup \mathbb{N}$. By (3.7) and (3.8), it holds that

$$\max_{\tilde{\sigma}_{1,\omega}} U_{1,\omega}(\tilde{\sigma}_{1,\omega}, \sigma^p_2) = \max_{\tilde{\sigma}_{1,\omega}} U_{1,\omega}(\tilde{\sigma}_{1,\omega}, \sigma_2) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2) = U_{1,\omega}(\sigma^p_{1,\omega}, \sigma_2) = U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2).$$

Therefore, public strategy $\sigma^p_{1,\omega}$ is the best response to σ^p_2.

Next, we show that $\sigma^p_2(h^t)$ is a player 2's best response for any $h^t \in \mathcal{H}$. By (3.7) and (3.8), for any $t \in \{0\} \cup \mathbb{N}$, $h^t \in \mathcal{H}^t$, $x \in [0, 1]$ and $a_2^t \in A_2$, it holds that

$$E\left[u_2(\cdot, a_2^t)|\sigma^p_1, \sigma^p_2, h^t, x\right] = E\left[u_2(\cdot, a_2^t)|\sigma_1, \sigma^p_2, h^t, x\right] = E\left[u_2(\cdot, a_2^t)|\sigma_1, \sigma_2, h^t, x\right].$$

We obtain that σ^p_2 is the best response to σ^p_1 because σ_2 is the best response to σ_1. Then, the strategy profile σ^p is a PE and $U_{1,\omega}(\sigma^p_{1,\omega}, \sigma^p_2) = U_{1,\omega}(\sigma_{1,\omega}, \sigma_2)$ for any $\omega \in \Omega$. Proposition 3.1 has been proved.

Chapter 4

Reputation and limited memory lead cooperative relationship

4.1 Introduction

We analyze reputation models: that is, infinitely repeated games with a long-run player, who has private information, and a sequence of short-run players. In most studies that analyze such infinitely repeated games, players are assumed to have *unlimited memories*. That is, short-run players can store all the information that they observed in the past. On the other hand, we assume that short- run players have *limited memories*. That is, each short-run player can observe signals in the fixed number of previous periods. We focus on equilibria that satisfy a certain condition, and compare the equilibria under the assumption of limited memories with those under the standard assumption of unlimited memories.

To be more specific, we consider infinitely repeated games with imperfect public monitoring in which each short run player can observe only public signals in the fixed number of previous periods. We consider a normal type and a bad type as long-run player's types. A normal type can choose his action without any restriction in each period. A *Bad type* commits to the stage-game Nash equilibrium action.

The assumption of limited memories has two effects on reputation. One is that reputation may be more fragile than under the assumption of unlimited memories. This is because even if relatively many good signals were realized in the past, when a short-run player has limited memory and has observed a few bad signals in recent periods, she believes that the long-run player is a bad type. The other effect is that reputation is recovered more easily than under the assumption of unlimited memories. This is because, under the assumption of limited memories, if a short-run player observes relatively few good signals, then she may believe that the long-run player is a normal type even when many bad signals were realized in the past.

We focus on the following feature of reputation. Imagine a relationship between a firm and consumers. When consumers believe that the firm produces a good product with similar probabilities, they choose similar actions. Conjecturing such short-run player's behavior, the firm also chooses similar actions when short-run players believe that the firm produces a good product with similar probabilities. That is, players choose similar actions when short-run players have similar beliefs.

To capture the feature of reputation, we focus on a class of equilibrium strategies: each player chooses a similar mixed action when a short-run player in a period has a similar belief as a short-run player in another period. This strategies are more restrictive than Markov strategies shown in Mailath and Samuelson (2001): if a short-run player in a period has the same belief as a short-run player in another period, then the players choose the same mixed actions. Our strategy captures the above feature of reputation, although Markov strategy cannot because it does not restrict players' action when short-run players have different beliefs.

An equilibrium is said to be *trivial* if a stage-game Nash equilibrium action profile is chosen in each period on the equilibrium path. If short-run players have unlimited memories, then they keep track infinitely long sequences of public signals. Hence, information contained in recent public signals becomes relatively small to the total information contained in the sequence as time goes by. Thus, players do not change their behavior at infinity because the beliefs of short-run players do not change depending on the recent signals. Therefore, the long-run player has a strong incentive to choose a dominant action of the stage-game because continuation play does not affected by recent public signals. This is a trivial equilibrium.

In contrast to the case of unlimited memory, if short-run players have limited memories, information of recent public signals might not become small at infinity. This is because, short-run player can store only finite public signals. In such a situation, short-run players change their actions depending on recent public signals, and it generates an incentive for the long-run player to choose an dominated action of the stage-game. We show that there exists a nontrivial equilibrium in which the long-run player chooses a dominated action of the stage-game to keep his reputation.

Typical models in the literature on reputation (e.g., Fudenberg and Levine (1989, 1992), Cripps et al. (2004, 2007)) assume unlimited memories. In some cases, however, the standard assumption of unlimited memories may not be plausible. Consider an online shop and its web-based customer reviews. There are typically many customer reviews for each product. However, online shops show only finite reviews in top page of each product page. (For example, Amazon.com shows six reviews on each product page.) Hence, many customers decide whether or not to buy the product based on only a finite number of reviews. In such a situation, the assumption of limited memories is more plausible.

There exist some previous studies about reputation under the assumption of limited memories: Liu (2011), Monte (2013) and Liu and Skrzypacz (2014). They consider infinitely repeated games with perfect monitoring. That is, players can observe the action chosen by the opponent player without monitoring error. They assume *Stackelberg type* as a commitment type. A Stackelberg type commits to an action that the long-run player is most likely to commit to. On the other hand, we consider a bad type as a commitment type.

Liu (2011) considers costly information acquisition. In his model, if short-run players pay a cost, then they can observe the long-run player's actions in the previous periods whose length depends on their payments. He finds a unique perfect Bayesian equilibrium in which short-run players observe public signal in the finite numbers of previous periods, and characterizes it. Thus, we can interpret the model in Liu (2011) as an endogenous limited memory model.

Liu and Skrzypacz (2014) consider a limited memory model. They show an equilibrium in which (i) the long-run player mimics a Stackelberg type with a positive probability to build a good reputation if short-run players doubt that he is a normal type, and (ii) he chooses a strictly dominant action in the stage-game if short-run players strongly believe that he is a Stackelberg type. On the other hand, we show an equilibrium in which the long- run player does not choose the strictly dominant action of the stage-game at any history.

The above two papers assume that short-run players cannot observe calendar time. When short-run players can observe calendar time, their beliefs change in complicated ways depending on calendar time and the analyses are difficult. We and Monte (2013) consider a limited memory model in which short-run players observe calendar time. Monte (2013) analyzes infinitely repeated zero-sum games in which short-run players can observe calendar time. He studies the beliefs of short-run players in the equilibrium. He shows that short-run players never know a long-run player's true type in any equilibrium. We also assume that short-run players can observe calendar time and show a nontrivial sequential equilibrium, although Monte (2013) shows the properties of short-run players' beliefs instead of the nontrivial equilibrium itself.

There exist only a few models that assume a bad type as a commitment type: Mailath and Samuelson (2001), Ely and Välimäki (2003) and Ely et al. (2008). Mailath and Samuelson (2001) assume unlimited memory and an *impermanent type*, whose type is replaced with a positive probability in each period. On the other hand, we consider limited memory and *permanent type*, whose type does not change throughout a play. Mailath and Samuelson (2001) focus on a Markov strategy in which the state space is the short-run player's belief about the long-run player's type. They show an equilibrium in which a normal type chooses a dominated action of the stage-game in each period in order to change the short-run player's belief.

Ely and Välimäki (2003) conducted a seminal work on the *bad reputation game*. Ely et al. (2008) generalize Ely and Välimäki (2003) and consider games in which short-run players have an option about whether to deal with the long-run player who can choose a *deceiving action* to look good in the current period. Ely et al. (2008) call the generalized games *bad reputation games*. Our model is not a bad reputation game because the stage-game in this chapter has no deceiving actions. We restrict the player's equilibrium strategy and find that the unique equilibrium is the repetition of the stage-game Nash equilibrium.

The rest of this chapter is organized as follows. We introduce the model in Section 4.2. Section 4.3 studies reputation under unlimited memory and Section 4.4 studies reputation under limited memory. Section 4.5 concludes.

4.2 Model

We consider the following stage-game. There are two players, player 1 and player 2. Let $A_i \equiv \{0, 1, \ldots, k_i\}$ be the finite set of actions. Let $u_i(a_1, a_2)$ be the expected stage-game payoff of player i given action profile $(a_1, a_2) \in A_1 \times A_2$. We assume that the expected stage-game payoffs satisfy the following assumptions.[1]

Assumption 4.1 For any $a_2 \in A_2$, the payoff of player 1, $u_1(a_1, a_2)$, is strictly decreasing in $a_1 \in A_1$ and, for any a_1, $u_1(a_1, a_2)$, is increasing in a_2.

That is, Assumption 4.1 implies that, in the stage-game, player 1 has a dominant action 0. In addition, for any a_1, $u_1(a_1, a_2)$ is the greatest when $a_2 = k_2$ holds.

Assumption 4.2 Fix any distribution functions on player 1's action F and F'. If F has first-order stochastic dominance over F', then the maximal element of player 2's

[1]The product-choice game defined in Mailath and Samuelson (2006) satisfies these Assumptions.

best reply to F is not smaller than the maximal element of player 2's best reply to F'. For any pure action $a_1 \in A_1$, player 2 has the unique best reply $a_2^*(a_1)$, such that $a_2^*(k_1) = k_2$ and $a_2^*(0) = 0$.

Assumption 4.2 implies that if player 1's mixed actions are ordered in the sense of first-order stochastic dominance, then player 2's best responses to the mixed actions are monotonic. In addition, if player 1 chooses pure action, then player 2 has a unique best reply. Assumption 4.1 and Assumption 4.2 imply that the stage-game has a unique Nash equilibrium.

Players cannot observe directly the opponent player's action, but, they can observe the same noisy signals. When an action profile (a_1, a_2) is chosen, a public signal y is realized with a probability $\rho(y|a_1, a_2)$ and observed by both players. Let a set $Y = \{y_0, y_2, \ldots, y_k\} \subset \mathbb{R}$ be the set of all public signals we assume that $y_s < y_{s+1}$ holds for any natural number $s(< k)$. We make the following assumptions on the distribution of public signals ρ.

Assumption 4.3 (Probability distribution independence) For any $a_1 \in A_1$, there exists $p_{a_1} \in (0,1)$ such that

$$\rho(y|a_1, a_2) = p_{a_1}(y) \in (0,1), \quad \forall a_2 \in A_2.$$

Assumption 4.3 ensures that for any action profile $a \in A$, any public signal y is realized with positive probability (full support assumption). Assumption 4.3 may seem to be restrictive. However, the following situations are conceivable. Consider an Amazon review. Suppose that several consumers bought a book from Amazon. It takes a long time to finish reading the book and they forget how much they paid for the book when they review the book. In such a situation, this assumption is plausible.

Assumption 4.4 (Monotone likelihood ratio property) If $a_1 < a_1'$, then $\frac{p_{a_1}(y)}{p_{a_1'}(y)}$ is a decreasing function on y.

Let $\pi_i(a_i, y)$ be the player i's stage-game payoff from an action a_i and a public signal y. Player i's payoff depends only on what player i knows because each player can observe only his own action and a public signal. Therefore, this payoff $\pi_i(a_i, y)$ provides no additional information about the action chosen by the opponent player. Then, we let $u_i(a_1, a_2) = \sum_{y \in Y} \pi_i(a_i, y)\rho(y|a_1, a_2)$ denote the ex ante stage payoff for player i.

We consider an infinitely repeated game where the stage-game is defined above. Players play the stage-game in period $t = 0, 1, \ldots$ under incomplete information. Player 1 is a long-run player with a discount factor $\delta \in [0, 1)$. Player 2 is a succession of short-run players who live for one period. We suppose that player 1 has a type. Let us denote a normal type by ω^* and a bad type by ω_0. Let $\Omega \equiv \{\omega^*, \omega_0\}$ be the set of all possible types. At the beginning of the infinitely repeated game, Nature selects player 1's type according to a common prior. With a probability $\mu \in (0, 1)$, player 1 is a normal type. With a probability $1 - \mu$, player 1 is a bad type. A normal type can choose an action from A_1 in every period. On the other hand, a bad type chooses action 0 in every period. Player 1's type is his private information and this does not change throughout a play of the infinitely repeated game.

Player 1 can observe a sequence of actions chosen by player 1 and public signals. Player 1 cannot observe player 2's actions.[2] For $t \geq 1$, we let $h_1^t = (a_1^s, y^s)_{s=0}^{t-1} \in (A_1 \times Y)^t$ denote the private history of player 1 at the beginning of period t.

Player 2 can observe only a sequence of public signals in a fixed number of periods. It is said that player 2 has *m-memory* if, for any $t \geq 1$, the private history of player 2 at period t, h_2^t, is a sequence of realized public signals from period $\max\{0, t - m\}$ to $t - 1$: that is, $h_2^t = (y^s)_{s=\max\{0,t-m\}}^{t-1} \in Y^{\min\{t,m\}}$. The natural number m is referred as *memory size*. If memory size is finite, then player 2 is said to have *limited memory*, and if memory size is infinite, then player 2 is said to have *unlimited memory*. If player 2 has m-memory, then for any $t \geq 1$, a private history of player 2, h_2^t, is common knowledge between player 1 and player 2 in period t. Therefore, for any

[2] This assumption is for simplicity of exposition. Even if player 1 can observe player 2's action, our results do not change.

$t \geq 1$, we regard private history h_2^t as a public history. For any private history of player 1, $h_1^t = (a_1^s, y^s)_{s=0}^{t-1} \in (A_1 \times Y)^t$, let us define $z(h_1^t)$ as the public history induced by h_1^t. That is, $z(h_1^t) = (y^s)_{s=\max\{0,t-m\}}^{t-1} = h_2^t$.

Given memory size m, let $\mathcal{H}_i^t$ be the set of all possible player i's private histories in period $t (\geq 1)$ and let $\mathcal{H}_i$ be the set of all possible player i's private histories. For any i, let $\mathcal{H}_i^0$ be an arbitrary singleton set.

Player 1's strategy σ_1 is a function that assigns a probability distribution over A_1 to each pair of type and private history of player 1. A bad type is assumed to commit to action 0. The strategy of player 2 in period t, σ_2^t, is a function that assigns a probability distribution over A_2 to each private history of player 2. Let us define player 2's strategy by $\sigma_2 \equiv (\sigma_2^t)_{t=0}^{\infty}$. We denote the strategy profile by $\sigma = (\sigma_1, \sigma_2)$.

Given a sequence of action profiles $(a_1^t, a_2^t)_{t=0}^{\infty}$, the average discounted payoff to a normal type with a discount factor $\delta \in [0, 1)$ is given by

$$(1-\delta) \sum_{t=0}^{\infty} \delta^t u_1(a_1^t, a_2^t).$$

A normal type maximizes the expected average discounted stage-game payoff given strategy of player 2. Player 2 in period t maximizes the expected stage-game payoff in period t.

A strategy profile $\sigma = (\sigma_1, \sigma_2)$ and a common prior μ produce a probability distribution over the set of player 1's types and all sequences of private histories. Let $\zeta_2^t(\omega, (h_1^s, h_2^s)_{s=0}^{t-1} | h_2^t, \sigma)$ be player 2's belief about player 1's type and private histories given player 2's private history h_2^t and strategy profile σ. Let $\zeta_1^t((h_2^s)_{s=0}^{t-1} | \omega, h_1^t, \sigma)$ be player 1's belief about player 2s' private histories given player 1's type, private history h_1^t and σ. We let $\zeta_i \equiv (\zeta_i^t)_{t=0}^{\infty}$ denote a system of beliefs of player i, and we define $\zeta \equiv (\zeta_1, \zeta_2)$. We do not explicitly refer to player 1's beliefs because they are determined uniquely by Bayes' rule.

Let $h_1^t \circ h_1^\tau$ be the concatenation of player 1's private history h_1^t followed by player 1's private history h_1^τ.

Definition 4.1 For any strategy of player 1, σ_1, player 1's strategy $\sigma_1|_{h_1^t}$ is *player 1's continuation strategy induced by* h_1^t, if strategy $\sigma_1|_{h_1^t}$ satisfies that

$$\sigma_1|_{h_1^t}(\omega, h_1^\tau) = \sigma_1(\omega, h_1^t \circ h_1^\tau), \quad \forall \omega \in \Omega, \ \forall h_1^\tau \in \mathcal{H}_1.$$

A strategy profile is *completely mixed* if every action profile is selected with positive probabilities at any history. Given a strategy profile σ, a system of beliefs is *consistent* if there exists a sequence of completely mixed strategy profiles converging to σ such that the corresponding sequence of the system of beliefs, obtained from Bayes' rule, converges to it.

We use sequential equilibrium as a solution concept.

Definition 4.2 A strategy profile σ is a *sequential equilibrium* if there exists a belief system ζ satisfying the following two conditions.

1. *Sequential Rationality*: For any t and for any h_1^t, player 1's continuation strategy induced by h_1^t, $\sigma_1|_{h_1^t}$, is a best response given ζ. For any t and for any h_2^t, the strategy of player 2 in period t, $\sigma_2^t(h_2^t)$, is a best response given ζ.

2. *Consistency of Belief System*: A belief system ζ is consistent with strategy profile σ.

In the next assumption, we define a class of strategies on which we focus. Let us denote $\eta_2^t(\omega|h_2^t)$ as player 2's belief about player 1's type given a player 2's private history and a strategy profile σ where we suppress σ from η_2^t.

We focus on a strategy profile in which arbitrary small changes in reputation do not affect players' behaviors in the following sense.

Assumption 4.5 For any $\gamma > 0$, there exists $\beta > 0$, such that, for any $h_1^t, h_1^{t'} \in \mathcal{H}_1$, it holds that [3]

$$|\eta_2^t(\omega^*|z(h_1^t)) - \eta_2^{t'}(\omega^*|z(h_1^{t'}))| < \beta \quad \Rightarrow \quad \begin{cases} ||\sigma_1(\omega^*, h_1^t) - \sigma_1(\omega^*, h_1^{t'})|| < \gamma, \\ ||\sigma_2(z(h_1^t)) - \sigma_2(z(h_1^{t'}))|| < \gamma. \end{cases}$$

[3] $||x||$ denotes the max norm of x.

An interpretation of Assumption 4.5 is as follows. Let us consider a firm (a long-run player) and consumers (short-run players). Consumers do not change their buying behavior depending on arbitrary small changes in reputation. If a firm has good reputation, then consumers continue to buy a product from the firm even if they buy a bad quality one only once.

Assumption 4.5 is more restrictive than the Markov strategy assumption in Mailath and Samuelson (2001). The Markov strategy restricts players' behaviors only when short-run players have the same belief, and does not restrict players' behaviors if short-run players have different beliefs.

In what follows, a strategy profile σ is said to be an *equilibrium* if it is a sequential equilibrium and satisfies Assumption 4.5. We call the equilibrium trivial if players play a stage-game Nash equilibrium on the equilibrium path in each period. An equilibrium is *nontrivial* if it is not trivial.

4.3 Unlimited memories

In this section, we analyze a model with unlimited memory as a benchmark. We show that if short-run players have unlimited memories, a unique equilibrium is a repetition of stage-game Nash equilibrium action profile $(0,0)$. In the following, we show the sketch of the proof of the statement by contradiction.

Let us consider a situation where short-run players have unlimited memories. Fix any nontrivial equilibrium σ. We define $\hat{\mathcal{H}}_1^t$ and $\overline{\eta}_2$ as follows:

$$\hat{\mathcal{H}}_1^t \equiv \left\{h_1^t \in \mathcal{H}_1^t \mid \sigma_1(h_1^t)(0) < 1 \text{ and } h_1^t \text{ is on the equilibrium path.}\right\},$$
$$\overline{\eta}_2 \equiv \sup\left\{x \in [0,1] \,\middle|\, x = \eta_2^t(\omega^*|z(h_1^t)) \quad \text{for some } t \text{ and } \hat{h}_1^t \in \hat{\mathcal{H}}_1^t\right\}.$$

There exist two cases for $\overline{\eta}_2$. One is the case of $\overline{\eta}_2 < 1$ and the other is the case of $\overline{\eta}_2 = 1$. We explain the first case. The second case is discussed in the proof. Fix $\overline{\eta}_2 (< 1)$ and suppose a private history $h_1^t \in \hat{\mathcal{H}}_1^t$ that satisfies $\eta_2^t(\omega^*|z(h_1^t)) = \overline{\eta}_2$. By the definition of $\hat{\mathcal{H}}_1^t$, it holds that $\eta_2^t(\omega^*|z(h_1^t) \circ y_k) > \overline{\eta}_2$ and this contradicts the definition of $\overline{\eta}_2$. Thus, there exists no private history h_1^t that satisfies $\eta_2^t(\omega^*|z(h_1^t)) = \overline{\eta}_2$, and for any $\varepsilon > 0$, there exists h_1^t such that $\overline{\eta}_2 - \eta_2^t(\omega^*|z(h_1^t)) < \varepsilon$.

Fix $\beta > 0$ and $\varepsilon \in (0, \beta)$. First, we fix $\alpha \in (0, 1)$ such that $(1-\alpha)(p_0(y_0)-p_1(y_0)) < \varepsilon$. Next, we fix $\varepsilon' > 0$, such that

$$\overline{\eta}_2 - \frac{(\overline{\eta}_2 - \varepsilon')(\alpha p_0(y_k) + (1-\alpha)p_1(y_k))}{(\overline{\eta}_2 - \varepsilon')(\alpha p_0(y_k) + (1-\alpha)p_1(y_k)) + (1 - (\overline{\eta}_2 - \varepsilon'))p_0(y_k)} < 0.$$

Consider a private history $h_1^t \in \hat{\mathcal{H}}_1^t$, such that $\overline{\eta}_2 - \eta_2^t(\omega^*|z(h_1^t)) < \varepsilon'$. If player 1 does not choose action 0 with more than probability α at the public history $z(h_1^t)$, then it holds that

$$\begin{aligned}\overline{\eta}_2 - \eta_2^{t+1}(\omega^*|z(h_1^t) \circ y_k) <& \overline{\eta}_2 - \frac{\eta_2^t(\omega^*|z(h_1^t))(\alpha p_0(y_k) + (1-\alpha)p_1(y_k))}{\eta_2^t(\omega^*|z(h_1^t))(\alpha p_0(y_k) + (1-\alpha)p_1(y_k)) + (1 - \eta_2^t(\omega^*|z(h_1^t)))p_0(y_k)} \\ <& 0\end{aligned}$$

That is, $\eta_2^{t+1}(\omega^*|z(h_1^t) \circ y_k) > \overline{\eta}_2$. It is a contradiction of the definition of $\overline{\eta}_2$. Thus, normal type chooses action 0 with at least probability α at public history $z(h_1^t)$. Then, it holds that

$$\begin{aligned}&\eta_2^t(\omega^*|z(h_1^t)) - \eta_2^t(\omega^*|z(h_1^t) \circ y_0) \\ <& \frac{\eta_2^t(\omega^*|z(h_1^t))(1 - \eta_2^t(\omega^*|z(h_1^t)))}{\eta_2^t(\omega^*|z(h_1^t)) \left\{\alpha p_0(y_0) + (1-\alpha)p_1(y_0)\right\} + (1 - \eta_2^t(\omega^*|z(h_1^t)))p_0(y_0)}(1-\alpha)(p_0(y_0) - p_1(y_0)) \\ <& \varepsilon\end{aligned}$$

Hence, we have $|\eta_2^t(\omega^*|z(h_1^t) \circ y) - \eta_2^t(\omega^*|z(h_1^t))| < \varepsilon (< \beta)$ for any y. It implies that $\overline{\eta}_2 - \eta_2^{t+1}(\omega^*|z(h_1^t) \circ y) < 2\beta$ and we can use the similar argument holds for public history $z(h_1^t) \circ y$.

For any $\beta > 0$, $\varepsilon > 0$ and k, we consider the same argument to public history $h_1^t \in \hat{\mathcal{H}}_1^t$ and $(y^s)_{s=1}^k$. Using the above discussion repeatedly, we have $\overline{\eta}_2 - \eta_2^{t+k}(\omega^*|z(h_1^t) \circ (y^s)_{s=1}^k) < 2^k\beta$ for any k and $(y^s)_{s=1}^k$. Hence, for any $k', k'' \leq k$, $(\tilde{y}^s)_{s=1}^{k'}$ and $(\hat{y}^s)_{s=1}^{k''}$, we have $|\eta_2^{t+k'}(\omega^*|z(h_1^t) \circ (\tilde{y}^s)_{s=1}^{k'}) - \eta_2^{t+k''}(\omega^*|z(h_1^t) \circ (\hat{y}^s)_{s=1}^{k''})| < 2^k\beta$. If β converges to zero, then players chooses the same mixed actions for any $h_1^t \circ (y^s)_{s=1}^k$ by Assumption 4.5.

That is, for any h_1^t, the continuation strategy profile induced by $z(h_1^t) \circ y$ does not change depending on y at least k periods. Therefore, normal type has a strong incentive to choose the dominant-action of the stage-game 0 at h_1^t in order to maximize the stage-game payoff. It is a contradiction to $h_1^t \in \hat{\mathcal{H}}_1^t$. We find that σ is not nontrivial.

Proposition 4.1 *Suppose that Assumption 4.1-4.5 hold and short run players have limited memories For any discount factor* $\delta \in [0, 1)$, *there exists no nontrivial equilibrium.*

Proof. See Appendix B.1. □

4.4 Limited memories

If short-run players have limited memories, short-run player cannot observe old signals. Hence, information of recent public signals does not become small as time goes by. Thus, short-run players might change their actions depending on the previous signals. Then, the long-run player has an incentive to choose action k_1 to keep his reputation.

In this section, we show that there exists an equilibrium in which the long-run player does not choose the dominant-action of the stage-game to keep his reputation under the assumption of limited memories. To be more specific, we show that there exists an equilibrium in which the normal type chooses action k_1 at any history.

We also impose the following assumption.

Assumption 4.6 (**Monotonicity of payoff**) If $a_1 > a_1'$ holds, then it holds that $u_1(a_1, a_2^*(a_1)) > u_1(a_1', a_2^*(a_1'))$.

This assumption determines which actions player 1 would most like to publicly commit. The above assumption ensures that player 1 prefers the largest action if he could publicly commit to.

Let $F_{a_1}(y)$ be a distribution function on Y, given a_1. The following proposition holds.

Proposition 4.2 *Suppose that Assumptions 4.1-4.6 hold and short-run players have m-memories. Assume that the following conditions* (4.1)–(4.3) *are satisfied:*

1.

$$k_2 \in \arg\max_{a_2} \frac{\mu p_{k_1}(y_k)^m u_2(k_1, a_2) + (1-\mu) p_0(y_k)^m u_2(0, a_2)}{\mu p_{k_1}(y_k)^m + (1-\mu) p_0(y_k)^m}, \tag{4.1}$$

$$0 \in \arg\max_{a_2} \frac{\mu p_{k_1}(y_0)^m u_2(k_1, a_2) + (1-\mu) p_0(y_0)^m u_2(0, a_2)}{\mu p_{k_1}(y_0)^m + (1-\mu) p_0(y_0)^m}, \tag{4.2}$$

2.

$$\max_{a_2 \in A_2} \{u_1(0, a_2) - u_1(k_1, a_2)\} < \delta^{m+1} \left\{ \min_{y \in Y} p_{k_1}(y) \right\}^{m-1} \{u_1(k_1, 1) - u_1(k_1, 0)\} \times \left[\min_{a_1 \in A_1 \setminus \{k_1\}} \min_{y \in Y \setminus \{y_k\}} \{F_{a_1}(y) - F_{k_1}(y)\} \right] \tag{4.3}$$

Then, there exists an equilibrium in which, player 1 *chooses action* k_1 *at any private history* h_1^t*.*

Proof. See Appendix B.2 □

We explain Proposition 4.2 by example. Suppose that $m = 1$, $A_1 = A_2 = \{0, 1\}$ and $Y = \{y_0, y_1\}$. Then, we can rewrite conditions (4.1)–(4.3) in Proposition 4.2 as follows.

$$1 \in \arg\max_{a_2} \frac{\mu p_1(y_1) u_2(1, a_2) + (1-\mu) p_0(y_1) u_2(0, a_2)}{\mu p_1(y_1) + (1-\mu) p_0(y_1)}, \tag{4.4}$$

$$0 \in \arg\max_{a_2} \frac{\mu p_1(y_0) u_2(1, a_2) + (1-\mu) p_0(y_0) u_2(0, a_2)}{\mu p_1(y_0) + (1-\mu) p_0(y_0)}, \tag{4.5}$$

$$\max_{a_2 \in A_2} \{u_1(0, a_2) - u_1(1, a_2)\} < \delta^2 \{u_1(1, 1) - u_1(1, 0)\} \times (p_1(y_1) - p_0(y_1)). \tag{4.6}$$

Hence, Condition 1 in Proposition 4.2 ensures that the short-run players' best response is 1 if the previous public signal is y_1, and it is 0 if the previous public signal is y_0. That is, the short-run players' best responses change depending on the previous public signals. Condition 2 in Proposition 4.2 ensures that player 1's deviation to action 0 from action 1 is not profitable.

We examine whether there exists an equilibrium profile σ in which, at any private history h_1^t, player 1 chooses action 1, or not. Fix σ_1 such that at any private history h_1^t, player 1 chooses action 1. Fix σ_2 such that player 2 chooses action 1 if and only if the previous public signal is y_1. We show that (σ_1, σ_2) is an equilibrium.

Let $v_1(y)$ be the continuation payoff of the normal type given a public signal y. The continuation payoff $v_1(y_i)$ is given by

$$v_1(y_i) =(1-\delta)u_1(1,i) + \delta p_1(y_1)v_1(y_1) + \delta(1-p_1(y_1))v(y_0).$$

Let $v_1^d(y_i)$ be the payoff when the previous public signal is y_i and player 1 deviates to action 0.

$$v_1^d(y_i) =(1-\delta)u_1(0,i) + \delta p_0(y_1)v_1(y_1) + \delta(1-p_0(y_1))v_1(y_0)$$

Hence, for any y_i, it holds that

$$\begin{aligned}
&v_1(y_i) - v_1^d(y_i)\\
=&(1-\delta)\{u_1(1,i) - u_1(0,i)\} + \delta(p_1(y_1) - p_0(y_1))(v_1(y_1) - v_1(y_0))\\
=&-(1-\delta)\{u_1(0,i) - u_1(1,i)\} + (1-\delta)\delta(p_1(y_1) - p_0(y_1))\{u_1(1,1) - u_1(1,0)\}\\
>&(1-\delta)\left[-\{u_1(0,i) - u_1(1,i)\} + \max_{a_2\in A_2}\{u_1(0,a_2) - u_1(1,a_2)\}\right]\\
\geq&0.
\end{aligned}$$

The strict inequality comes from Condition 2. Hence, σ_1 is a best response to σ_2. Condition 1 ensures that σ_2 is a best response to σ_1. Therefore, the strategy (σ_1, σ_2) is a sequential equilibrium.

Next, we show that strategy (σ_1, σ_2) satisfies Assumption 4.5. A normal type chooses action 1 at any history. Hence, if previous signal is y_1 (resp. y_0), then short-run player has a belief $\frac{\mu p_1(y_1)}{\mu p_1(y_1)+(1-\mu)p_0(y_1)}$ (resp. $\frac{\mu p_1(y_0)}{\mu p_1(y_0)+(1-\mu)p_0(y_0)}$). Thus, the following three kinds of beliefs are realized on the equilibrium path.

$$\frac{\mu p_1(y_1)}{\mu p_1(y_1)+(1-\mu)p_0(y_1)}, \quad \mu, \quad \frac{\mu p_1(y_0)}{\mu p_1(y_0)+(1-\mu)p_0(y_0)}.$$

The second one is a common prior.

We define β as follows.

$$\beta \equiv \min\left\{\frac{\mu p_1(y_1)}{\mu p_1(y_0)+(1-\mu)p_0(y_0)}-\mu, \mu-\frac{\mu p_1(y_0)}{\mu p_1(y_1)+(1-\mu)p_0(y_1)}\right\}.$$

By the definition of β, for any h_1^t and $h_1^{t'}$, if $|\eta_2^t(\omega^*|z(h_1^t))-\eta_2^{t'}(\omega^*|z(h_1^{t'}))|<\beta$, then $z(h_1^t)=z(h_1^{t'})$ holds. Thus, for any $\gamma>0$, the strategy (σ_1,σ_2) satisfies Assumption 4.5.

Finally, we consider a relation between memory size and Conditions 1 and 2 in Proposition 4.2. Condition 1 is not satisfied if memory size is small. This is because if memory size is small, then short-run player can observe only a small number of public signals and the beliefs of short-run players do not change sufficiently enough to change their best response. Condition 2 is not satisfied if memory size is large. This is because, if many public signals y_0 are realized in the past and memory size is large, it takes a long time to recover his reputation. Then, it does not pay for player 1 to choose action k_1 in order to recover his reputation. Hence, we can use Proposition 4.2 if memory size is neither too small nor too large. Thus, players can maintain good relationshiops in the long-run for moderate memory size.

4.5 Conclusion

In section 4.3, we have shown that if short-run players have unlimited memories, the unique equilibrium is a repetition of a stage-game Nash equilibrium action profile, which is Pareto inefficient. That is, if short-run players have unlimited memories, then players do not choose cooperative behavior. In Section 4.4, we have shown that if short-run players have limited memories, then there exists a nontrivial equilibrium in which a normal type chooses action k_1 at any history. In other words, if short-run players have limited memories, then players choose cooperative behavior in some equilibrium. Limited memories sometimes increase the incentive of the long-run player to choose action k_1 at any history in order to keep his reputation, so that it is good not only for the normal type, but also for short-run players.

Appendix

B.1 Proof of Proposition 4.1

Proof of Proposition 4.1. We prove Proposition 4.1 by contradiction. Fix any non-trivial equilibrium σ. We fix an increasing sequence of natural numbers $(\ell_s)_{s=1}^{\infty}$ and a sequence of private histories $(\hat{h}_1^{\ell_s})_{s=1}^{\infty}$, such that

$$\begin{aligned}
&\lim_{s\to\infty} \eta_2^{\ell_s}(\omega^*|z(\hat{h}_1^{\ell_s})) = \overline{\eta}_2(>\mu),\\
&\eta_2^{\ell_s}(\omega^*|z(\hat{h}_1^{\ell_s})) < \overline{\eta}_2, \quad \forall s\in\mathbb{N}, \quad \text{and}\\
&\hat{h}_1^{\ell_s} \in \hat{\mathcal{H}}_1^{\ell_s}, \quad \forall s\in\mathbb{N}.
\end{aligned}$$

In what follows, we show that, there exists a natural number s^* for which a normal type's unique best response is 0 at private history $\hat{h}_1^{\ell_{s^*}}$.

First, we consider a case where $\overline{\eta}_2 = 1$ holds. We fix $\overline{\gamma}$ and $\overline{T}$, such that

$$\min_{a_2\in A_2}\{u_1(0,a_2)-u_1(1,a_2)\} > 2\frac{\delta(1-\delta^{\overline{T}})}{1-\delta}\overline{\gamma}\{u_1(0,k_2)-u_1(k_1,0)\} + \delta^{\overline{T}+1}\{u_1(0,k_2)-u_1(k_1,0)\}.$$

For any $\overline{\beta}(>0)$, we fix $\overline{s}$ that satisfies

$$\frac{\eta_2^{\ell_{\overline{s}}}(\omega^*|z(\hat{h}_1^{\ell_{\overline{s}}}))\{\rho(y_0|k_1)\}^{\overline{T}+1}}{\eta_2^{\ell_{\overline{s}}}(\omega^*|z(\hat{h}_1^{\ell_{\overline{s}}}))\{\rho(y_0|k_1)\}^{\overline{T}+1} + (1-\eta_2^{\ell_{\overline{s}}}(\omega^*|z(\hat{h}_1^{\ell_{\overline{s}}})))\{\rho(y_0|0)\}^{\overline{T}+1}} > 1-\overline{\beta}.$$

Then, it holds that

$$1-\eta_2^{(\ell_{\overline{s}})+t'}(\omega^*|z(\hat{h}_1^{\ell_{\overline{s}}}\circ h_1^{t'})) < \overline{\beta}, \quad \forall h_1^{t'}\in\mathcal{H}_1^{t'}, \quad \forall t'\in\{0,1,\ldots,\overline{T}+1\}. \qquad (4.7)$$

It ensures that

$$|\eta_2^{(\ell_{\overline{s}})}(\omega^*|z(\hat{h}_1^{\ell_{\overline{s}}}) - \eta_2^{(\ell_{\overline{s}})+t'}(\omega^*|z(\hat{h}_1^{\ell_{\overline{s}}} \circ h_1^{t'}))| < \overline{\beta}, \quad \forall h_1^{t'} \in \mathcal{H}_1^{t'}, \quad \forall t' \in \{0, 1, \ldots, \overline{T}+1\}. \tag{4.8}$$

Hence, it holds under Assumption 4.5 that there exists $\overline{\beta} > 0$ and the corresponding $\overline{s}$ such that for any $t' \in \{0, 1, \ldots, \overline{T}+1\}$, and for any $h_1^{t'} \in \mathcal{H}_1^{t'}$,

$$\begin{aligned}
&\left| u_1(\sigma_1(\omega^*, \hat{h}_1^{\ell_{\overline{s}}}), \sigma_2(z(\hat{h}_1^{\ell_{\overline{s}}}))) - u_1(\sigma_1(\omega^*, \hat{h}_1^{\ell_{\overline{s}}} \circ h_1^{t'}), \sigma_2(z(\hat{h}_1^{\ell_{\overline{s}}} \circ h_1^{t'}))) \right| \\
<&\overline{\gamma} \left\{ u_1(0, k_2) - u_1(k_1, 0) \right\}.
\end{aligned} \tag{4.9}$$

Let us denote by $v_1(a_1, h_1^t)$ the continuation payoff when the normal type chooses a one-shot deviation from σ_1 to an action a_1 at private history h_1^t. Based on inequality (4.9), the continuation payoffs $v_1(a_1, \hat{h}_1^{\ell_{\overline{s}}})$ and $v_1(0, \hat{h}_1^{\ell_{\overline{s}}})$ satisfy that

$$\begin{aligned}
v_1(a_1, \hat{h}_1^{\ell_{\overline{s}}}) <&(1-\delta) u_1(a_1, \sigma_2(z(h_1^{\ell_{\overline{s}}}))) \\
&+ \delta(1-\delta^{\overline{T}}) \left[u_1(\sigma_1(\omega^*, \hat{h}_1^{\ell_{\overline{s}}}), \sigma_2(z(\hat{h}_1^{\ell_{\overline{s}}}))) + \overline{\gamma} \left\{ u_1(0, k_2) - u_1(k_1, 0) \right\} \right] \\
&+ \delta^{\overline{T}+1}(1-\delta) u_1(0, k_2), \quad \forall a_1 \in A_1 \backslash \{0\}, \\
v_1(0, \hat{h}_1^{\ell_{\overline{s}}}) >&(1-\delta) u_1(0, \sigma_2(z(h_1^{\ell_{\overline{s}}}))) \\
&+ \delta(1-\delta^{\overline{T}}) \left[u_1(\sigma_1(\omega^*, \hat{h}_1^{\ell_{\overline{s}}}), \sigma_2(z(\hat{h}_1^{\ell_{\overline{s}}}))) - \overline{\gamma} \left\{ u_1(0, k_2) - u_1(k_1, 0) \right\} \right] \\
&+ \delta^{\overline{T}+1}(1-\delta) u_1(k_1, 0).
\end{aligned}$$

Therefore, for any $a_1 \in A_1 \backslash \{0\}$, it holds that

$$\begin{aligned}
&\frac{v_1(a_1, \hat{h}_1^{\ell_{\overline{s}}}) - v_1(0, \hat{h}_1^{\ell_{\overline{s}}})}{1-\delta} \\
<& - \left[\min_{a_2 \in A_2} \left\{ u_1(0, a_2) - u_1(1, a_2) \right\} \right] \\
&+ 2 \frac{\delta(1-\delta^{\overline{T}})}{1-\delta} \overline{\gamma} \left\{ u_1(0, k_2) - u_1(k_1, 0) \right\} + \delta^{\overline{T}+1} \left\{ u_1(0, k_2) - u_1(k_1, 0) \right\} \\
<&0.
\end{aligned}$$

It is proved that there exists a natural number $\overline{s}$ such that the normal type's unique best response is action 0 at private history $\hat{h}_1^{\ell^*_{\overline{s}}}$ when $\overline{\eta}_2 = 1$.

Next, we consider a case where $\overline{\eta}_2 < 1$ holds. Suppose that $\lim_{s\to\infty} \sigma_1(\omega^*, \hat{h}_1^{\ell_s})(0) \neq 1$. Fix a sufficiently large s. When a public signal y_k is realized given a private history $\hat{h}_1^{\ell_s}$, then the belief $\eta_2^{(\ell_s)+1}(\omega^*|z(\hat{h}_1^{\ell_s})\circ y_k)$ is greater than $\overline{\eta}_2$ because $\lim_{s\to\infty} \sigma_1(\omega^*, \hat{h}_1^{\ell_s})(0) \neq 1$ and $\lim_{s\to\infty} \eta_2^{\ell_s}(\omega^*|z(\hat{h}_1^{\ell_s})) = \overline{\eta}_2$. We find that player 1's strategy satisfies $\lim_{s\to\infty} \sigma_1(\omega^*, \hat{h}_1^{\ell_s})(0) = 1$. Hence, for any $T \in \mathbb{N}$, there exists $s \in \mathbb{N}$, such that

$$\sigma_2(z(\hat{h}_1^{\ell_s} \circ h_1^{t'}))(0) = 1, \quad \forall h_1^{t'} \in \mathcal{H}_1^{t'} \quad \forall t' \in \{0, 1, \ldots, T+1\}.$$

By Assumption 4.2, player 2's strategy satisfies $\lim_{s\to\infty} \sigma_2(z(\hat{h}_1^{\ell_s}))(0) = 1$.

We fix natural numbers $\underline{T}$ and $\underline{s} \in \mathbb{N}$ that satisfy

$$\begin{aligned} u_1(0,0) - u_1(1,0) > & \delta^{\underline{T}+1} \left\{ u_1(0,k_2) - u_1(k_1,0) \right\}, \\ \sigma_1(\omega^*, \hat{h}_1^{\ell_{\underline{s}}} \circ h_1^{t'})(0) > 0, & \quad \forall h_1^{t'} \in \mathcal{H}_1^{t'} \quad \forall t' \in \{0, 1, \ldots, \underline{T}+1\}, \quad \text{and} \\ \sigma_2(z(\hat{h}_1^{\ell_{\underline{s}}} \circ h_1^{t'}))(0) = 1, & \quad \forall h_1^{t'} \in \mathcal{H}_1^{t'} \quad \forall t' \in \{0, 1, \ldots, \underline{T}+1\}. \end{aligned}$$

Let us denote by $v_1(a_1, h_1^t)$ the continuation payoff when the normal type chooses one shot deviation from σ_1 to action a_1 at a private history h_1^t. Continuation payoffs $v_1(a_1, \hat{h}_1^{\ell_{\underline{s}}})$ and $v_1(0, \hat{h}_1^{\ell_{\underline{s}}})$ satisfy that

$$\begin{aligned} v_1(a_1, \hat{h}_1^{\ell_{\underline{s}}}) < & (1-\delta)u_1(a_1,0) + \delta(1-\delta^{\underline{T}})u_1(0,0) + \delta^{\underline{T}+1}(1-\delta)u_1(0,k_2), \quad \forall a_1 \in A_1\backslash\{0\}, \\ v_1(0, \hat{h}_1^{\ell_{\underline{s}}}) \geq & (1-\delta)u_1(0,0) + \delta(1-\delta^{\underline{T}})u_1(0,0) + \delta^{\underline{T}+1}(1-\delta)u_1(k_1,0). \end{aligned}$$

Thereby, for all $a_1 \in A_1\backslash\{0\}$, it holds that

$$\frac{v_1(a_1, \hat{h}_1^{\ell_{\underline{s}}}) - v_1(0, \hat{h}_1^{\ell_{\underline{s}}})}{1-\delta} < -\left(u_1(0,0) - u_1(1,0)\right) + \delta^{\underline{T}+1} \left\{ u_1(0,k_2) - u_1(k_1,0) \right\} < 0.$$

It is proved that there exists a natural number $\underline{s}$ such that the normal type's unique best response is action 0 when $\overline{\eta}_2 < 1$. Therefore, it has been proved that there exists a natural number s^*, such that the normal type's unique best response is action 0 at private history $\hat{h}_1^{\ell_{s^*}}$ when $\underline{\eta}_2 = 1$. Hence, σ is not an equilibrium. It is a contradiction.

□

B.2 Proof of Proposition 4.2

Proof of Proposition 4.2. Let us define σ_1 as the strategy of player 1 in which the normal type chooses k_1 at any history. Let us define σ_2 as the pure strategy of player 2 in which, at any h_2^t, player 2 chooses the largest action among the best responses given σ_1 and η_2. In what follows, we prove that the strategy profile (σ_1, σ_2) is a sequential equilibrium.

For any public history $h_2^t = (y^s)_{\max\{0,t-m\}}^{t-1}$, we define $\Pr(h_2^t)$ as follows:

$$\Pr(h_2^t) = \prod_{s=\max\{0,t-m\}}^{t-1} p_{k_1}(y^s).$$

The probability $\Pr(h_2^t)$ is equal to the probability that public history h_2^t is realized when player 1 chooses action k_1 in each period from period $\max\{0, t-m\}$ to period $t-1$.

For any t and the sequence of public signal $(y^s)_{s=0}^{t-1}$, let us define $\tilde{z}((y^s)_{s=0}^{t-1})$ as the public history induced by $(y^s)_{s=0}^{t-1}$. That is, $\tilde{z}((y^s)_{s=0}^{t-1}) = (y^s)_{s=\max\{0,t-m\}}^{t-1}$. For any h_2^t and $\tilde{h}_2^{t'}$, let us define $h_2^t \oplus \tilde{h}_2^{t'} \equiv \tilde{z}(h_2^t \circ \tilde{h}_2^{t'})$. That is, $h_2^t \oplus \tilde{h}_2^{t'}$ is the sequence of public signals from m-period ago given a concatenation of h_2^t followed by $\tilde{h}_2^{t'}$.

The distribution function on A_1 given belief at $(h_2^t \oplus y_s) \oplus h_2^{t'}$ has first-order stochastic dominance over that given belief at $(h_2^t \oplus y_{s-1}) \oplus h_2^{t'}$. Hence, by Assumption 4.2, for any s, it holds that

$$u_1(a_1, \sigma_2((h_2^t \oplus y_s) \oplus h_2^{t'})) \geq u_1(a_1, \sigma_2((h_2^t \oplus y_{s-1}) \oplus h_2^{t'})).$$

Let us denote by $v_1(a_1, h_2^t)$ the continuation payoff when the normal type chooses a one-shot deviation from σ_1 to action a_1 at public history h_2^t. The continuation payoff

$v_1(a_1, h_2^t)$ is given by

$$
\begin{aligned}
v_1(a_1, h_2^t) =& (1-\delta)u_1(a_1, \sigma_2(h_2^t)) \\
&+ (1-\delta)\sum_{t'=0}^{\infty}\delta^{t'+1}\sum_{y\in Y}p_{a_1}(y)\sum_{h_2^{t'}\in\mathcal{H}_2^{t'}}\Pr(h_2^{t'})u_1(k_1, \sigma_2((h_2^t\oplus y)\oplus h_2^{t'})) \\
=& (1-\delta)u_1(a_1, \sigma_2(h_2^t)) \\
&+ (1-\delta)\sum_{t'=0}^{\infty}\delta^{t'+1}\sum_{h_2^{t'}\in\mathcal{H}_2^{t'}}\Pr(h_2^{t'})\Bigg[u_1(k_1, \sigma_2((h_2^t\oplus y_k)\oplus h_2^{t'})) \\
&+ \sum_{s=1}^{k}\{u_1(k_1, \sigma_2((h_2^t\oplus y_s)\oplus h_2^{t'})) - u_1(k_1, \sigma_2((h_2^t\oplus y_{s-1})\oplus h_2^{t'}))\}\{-F_{k_1}(y_{s-1})\}\Bigg].
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
v_1(k_1, h_2^t) - v_1(a_1, h_2^t) =& -(1-\delta)\{u_1(a_1, \sigma_2(h_2^t)) - u_1(k_1, \sigma_2(h_2^t))\} \\
&+ (1-\delta)\sum_{t'=0}^{\infty}\delta^{t'+1}\sum_{h_2^{t'}\in\mathcal{H}_2^{t'}}\Pr(h_2^{t'})\Bigg[\sum_{s=1}^{k}\{u_1(k_1, \sigma_2((h_2^t\oplus y_s)\oplus h_2^{t'})) \\
&- u_1(k_1, \sigma_2((h_2^t\oplus y_{s-1})\oplus h_2^{t'}))\}\{F_{a_1}(y_{s-1}) - F_{k_1}(y_{s-1})\}\Bigg].
\end{aligned} \tag{4.10}
$$

By first-order stochastic dominance, for any $h_2^t, h_2^{t'}$, it holds that

$$
\sum_{s=1}^{k}\left[\{u_1(k_1, \sigma_2((h_2^t\oplus y_s)\oplus h_2^{t'})) - u_1(k_1, \sigma_2((h_2^t\oplus y_{s-1})\oplus h_2^{t'}))\}\{F_{a_1}(y_{s-1}) - F_{k_1}(y_{s-1})\}\right] \geq 0.
$$

The difference in the expected stage-game payoff $(1-\delta)\{u_1(a_1, \sigma_2(h_2^t)) - u_1(k_1, \sigma_2(h_2^t))\}$ is positive. To ensure that the right-hand side of (4.10) is positive, we will show that there exist $h_2^t, h_2^{t'}$ for which the last term of (4.10) is positive.

For any $s \in \{0, 1, \ldots, m-1\}$, let us define g_s as

$$
g_s = (\overbrace{y_0, y_0, \ldots, y_0}^{m-s-1}, \overbrace{y_k, y_k, \ldots, y_k}^{s}).
$$

By (4.1) and (4.2), there exists a $(\xi, \psi)(\in \{0, 1, \ldots, k\} \times \{0, 1, \ldots, m-1\})$ such that

$$
\sigma_2(y_\xi \oplus g_\psi) > \sigma_2(y_{\xi-1} \oplus g_\psi) = 0.
$$

For any g_i and h_2^t, it holds that $(h_2^t \oplus y) \oplus g_i = y \oplus g_i$ and

$$u_1(k_1, \sigma_2((h_2^t \oplus y_\xi) \oplus g_\psi)) - u_1(k_1, \sigma_2((h_2^t \oplus y_{\xi-1}) \oplus g_\psi)) > u_1(k_1, 1) - u_1(k_1, 0).$$

Therefore, we obtain

$$\begin{aligned}
& v_1(k_1, h_2^t) - v_1(a_1, h_2^t) \\
\geq & -(1-\delta) \max_{a_2 \in A_2} \{u_1(0, a_2) - u_1(k_1, a_2)\} \\
& + (1-\delta)\delta^{m+1} \Pr(g_\psi) \{u_1(k_1, 1) - u_1(k_1, 0)\} \{F_{a_1}(y_{\xi-1}) - F_{k_1}(y_{\xi-1})\} \\
\geq & -(1-\delta) \max_{a_2 \in A_2} \{u_1(0, a_2) - u_1(k_1, a_2)\} + (1-\delta)\delta^{m+1} \left\{ \min_{y \in Y} p_{k_1}(y) \right\}^{m-1} \{u_1(k_1, 1) - u_1(k_1, 0)\} \\
& \times \left[\min_{a_1 \in A_1 \setminus \{k_1\}} \min_{y \in Y \setminus \{y_k\}} \{F_i(y) - F_{k_1}(y)\} \right] > 0.
\end{aligned}$$

It is proved that σ_1 is an optimal strategy for player 1. The strategy of player 2 is an optimal strategy of player 2 by the definition. Hence, it is proved that (σ_1, σ_2) is a sequential equilibrium.

Finally, we show that strategy profile (σ_1, σ_2) satisfies Assumption 4.5. Player 1's strategy satisfies Assumption 4.5 because the normal type chooses action k_1 at any history. Let us define a finite set B as the set of beliefs that are realized on the equilibrium path:

$$\begin{aligned}
B \equiv & \{x \in [0,1] | x = \eta_2^t(\omega^* | h_2^t) \text{ for some } t \text{ and } h_2^t\} \\
= & \{x \in [0,1] | x = \eta_2^t(\omega^* | h_2^t) \text{ for some } t \in \{0, 1 \ldots, m\} \text{ and } h_2^t \in \mathcal{H}_2^t\}
\end{aligned}$$

We select β such that for any $x, x' \in B$

$$||x - x'|| < \beta \Rightarrow x = x'.$$

By the definition of σ_2, player 2 chooses the same (mixed) actions if player 2 has the same beliefs. Hence, it is proved that player 2's strategy satisfies Assumption 4.5. The strategy profile (σ_1, σ_2) is an equilibrium. □

Chapter 5

An Efficiency in Repeated Prisoners' Dilemma with Observation Costs

5.1 Introduction

It is well known that if each player in an infinitely repeated game can observe the action of the other players, then they can achieve a Pareto efficient equilibrium payoff vector (the efficiency theorem). The efficiency theorem is the following statement: a Pareto efficient payoff vector is achieved by an equilibrium when players are patient. However, it is not obvious whether an efficiency theorem holds under private monitoring structures.

The monitoring structure is said to be costly observation if a player obtains additional information when he incurs a cost. We focus on a costly observation as a monitoring structure, and we show an efficiency result in an infinitely repeated prisoner's dilemma under costly observation.

A few papers present the folk theorem in some types of games under costly observation. The folk theorem is the following statement: Any feasible and individual rational payoff vector is achieved by an equilibrium if players are patient. Hence, if the folk theorem holds, then efficiency theorem holds. Ben-Porath and Kahneman (2003) show a folk theorem when communication is available. Miyagawa et al. (2003) show an approximate folk theorem when each player can choose at least three actions and the monitoring cost is small.

Our main contribution is to show an efficiency result in an infinitely repeated prisoner's dilemma, which the above results do not cover. The basic idea of the current chapter is public randomization (sunspot) and restart as shown in Bhaskar and van Damme (2002). In each period, a sunspot is realized at the end of the period. We consider the following strategy: for any period, when a specific sunspot is realized at the end of the period, the continuation strategy from the next period is equal to the strategy from initial period. Under such strategies, the continuation plays are determined independently of the realized signals if a specific sunspot is realized at the end of the period. Thus, each player views the current stage-game payoff as more important than when there exists no restart. Hence, it is well known that, under such strategies, players play a game as if they have lower discount factors than their own discount factors.

In particular, we confine our attention to a sequential equilibrium in which each player follows an automaton strategy. We say that a player follows an automaton strategy if he chooses an action and a monitoring decision depending on his private states. The key idea of the current chapter is that we change the restart probability depending on states. In previous studies, the restart probability is assumed to be the same for any history. However, we assume that the restart probability in the initial state is smaller than in the other states. This new idea helps players to coordinate with the other players, and we show that an efficiency result holds when the observation costs are sufficiently small in repeated games with less complex strategy than previous papers.

Many papers investigate whether a folk theorem or efficiency theorem holds or not in infinitely repeated games under various monitoring structures. The monitoring structure is said to be *perfect* if each player observes the realized action profile. Monitoring structure is said to be *imperfect public* if players cannot observe realized action profile, but they can observe the same noisy signals. The monitoring structure is said to be *(imperfect) private* if each player cannot observe the realized action profile, but he can observe a signal, which is realized stochastically and his private information.

A folk theorem is proved under perfect monitoring by Fudenberg and Maskin (1986), and is proved under imperfect public monitoring by Fudenberg et al. (1994). That is, they show that any feasible and strictly individual rational payoff vectors can achieve by an equilibrium under each monitoring structure.

There are various studies on private monitoring because there is a variety of private monitoring (e.g., conditional independent private signals, correlated private signals, and so on). Many papers try to prove a folk theorem or an efficiency result under each private monitoring structure. Ben-Porath and Kahneman (1996), Kandori and Matsushima (1998), Compte (1998) and Obara (2009) show folk theorems under private monitoring with communication. A private monitoring structure is said to be almost perfect if signals are sufficiently informative. Ely and Välimäki (2002) and Hörner and Olszewski (2006) show folk theorems without communication under almost perfect monitoring. Sugaya and Wolitzky (2014) show a folk theorem when a mediator is available.

Our model with costly observation belongs to a class of games with private monitoring. However, it is significantly different from the above studies on private monitoring. In costly observation, each player's monitoring decision is assumed to be unobservable. This structure makes it difficult to enforce each player to monitor the other player. If any player does not monitor the other player, each player strictly prefers the dominant-action of the stage-game. Thus, it is difficult to achieve a Pareto efficient equilibrium payoff vector under costly observation.

Some papers challenge to analyze infinitely repeated games under costly observation. Lehrer (1989, 1992a, 1992b) provides the seminal works about costly observation. He considers the following two actions. One action is dominated by the other action in the stage-game. These two actions produce the same probability distribution over the set of other player's signals. That is, the other player cannot distinguish between these two actions. He shows that, in an equilibrium, a player can choose the dominated action of the stage-game if he obtains precise information from the action. This can be interpreted as a player choosing costly monitoring in an equilibrium.

Ben-Porath and Kahneman (2003) and Miyagawa et al. (2003) consider a costly observation model in which (i) if a player pays a cost, then he observes the action chosen by other players, and (ii) if a player does not pay a cost, then he observes nothing. Miyagawa et al. (2003) show a sufficient condition for an efficiency result with public randomization when the monitoring cost is sufficiently small and each player is patient. Ben-Porath and Kahneman (2003) introduce communication and show a sufficient condition for a folk theorem. Miyagawa et al. (2008) relaxed the assumption in ours and the above two studies that each player observes nothing if he does not incur a cost. Miyagawa et al. (2008) assume that each player can observe private signals even if he does not incur a cost. They show a folk theorem in repeated games without communication for any level of observation costs.

Our monitoring structure is the same as those analyzed in the above three studies in the sense that if a player pays a cost, then he can observe the realized action profile in the current period. We show an efficiency result with less complex strategy (three-state automata) than those in the studies (e.g., six-state automata in a prisoner's dilemma in Miyagawa et al. (2008)).

Kandori and Obara (2004) assume that each player can observe not only the other player's action but also the other player's monitoring decision, if he incurs a cost. They allow monitoring error about "monitoring decisions". That is, when players monitor the other players, each player might observe a different signal from the monitoring decision chosen by the other player. They show an efficiency result when a monitoring error occurs with a small probability.

The following two studies relaxed the assumption in the above studies that if a player pays a cost, then he can observe the realized action profile in the "current" period. Flesch and Perea (2009) assume that each player can observe "actions in the past" if they incur costs. They show that if players can choose at least four actions, then a folk theorem holds even when neither public randomization nor communication is available. Our model discussed in this chapter does not satisfy this condition. Awaya (2014) checks the robustness of Takahashi (2010). Takahashi (2010) shows a folk theorem in repeated games with randomly matched players. Awaya (2014) assumes the following costly observation: each player can observe "the sequence of actions that the opposing player chose in the past" if he monitors the other player. He consider the monitoring cost is infinitesimal. That is, the monitoring decision does not affect each player's payoff, but he prefers a strategy in which the number of monitoring is small among strategies under which he obtains the same expected stage-game payoff (lexicographic preference). Awaya (2014) proves a folk theorem in models of Takahashi (2010) under infinitesimal monitoring cost when communication is available. He also shows that the strategy he uses to prove the folk theorem is not a sequential equilibrium if the observation cost is strictly positive. In addition, if the observation cost is greater than the maximum difference of stage-game payoffs, then any equilibrium is a repetition of stage-game Nash equilibrium action profile.

The rest of this chapter is organized as follows. We introduce the model in Section 5.2. In Section 5.3, we show an efficiency result.

5.2 Model

In this section, we define an infinitely repeated game with costly observation. That is, monitoring of each player is voluntary and it incurs a cost. First, we define a prisoner's dilemma as a stage-game and explain monitoring structure. Next, we consider the infinitely repeated prisoner's dilemma.

We consider a prisoner's dilemma as a stage-game. Let $A_i \equiv \{C_i, D_i\}$ be the set of actions available for player i. Let $A \equiv A_1 \times A_2$ denote the set of action profiles. Given an action profile $a \in A$, the stage-game payoff for player i, $u_i(a)$, is given by the following payoff matrix.

Prisoner's dilemma

Player 2

		C_2	D_2
Payer 1	C_1	$1, 1$	$-\ell, 1+g$
	D_1	$1+g, -\ell$	$0, 0$

Assumption 5.1 (i) $g > 0$ and $\ell > 0$, (ii) $g - \ell < 1$ and (iii) $g - \ell > 0$.

The first condition implies that C_i is dominated by D_i, and the second condition ensures that the payoff vector of action profile (C_1, C_2) is Pareto-efficient. The last condition is crucial for our result.

Each player decides which action to choose and whether to monitor the opponent or not, at the same time. If player i chooses to monitor, he incurs monitoring cost $\lambda > 0$ and he observes the action chosen by the other player. If player i does not monitor, he does not incur any additional cost. Let $B_i \equiv \{M_i, N_i\}$ be the set of monitoring decision of player i. Monitoring decision M_i means that player i chooses to monitor. Monitoring decision N_i means that player i chooses not to monitor. Hence, given an action profile $a \in A$ and monitoring decision $b_i \in B_i$, the stage-game payoff for player i, $\tilde{u}_i(a, b_i)$, is given by

$$\tilde{u}_i(a, b_i) \equiv \begin{cases} u_i(a) - \lambda, & \text{if } b_i = M_i, \\ u_i(a), & \text{if } b_i = N_i. \end{cases}$$

Let $O_i \equiv A_j \cup \{\varphi_i\}$ be the set of observations for player i. Observing $a_j \in A_j$ means that player i chooses monitoring decision M_i and observes that player j chooses action a_j. If player i chooses monitoring decision N_i, then he observes φ_i, that is, he observes nothing. We assume that monitoring decision of player i is not observable to player j. Each player cannot observe the monitoring decision of the opponent player even when he pays an observation cost.

Players play the above prisoner's dilemma repeatedly over periods $t = 1, 2, \ldots$ We assume that there exists a public randomization device ("sunspot"). In each period, a sunspot is realized after the choice of actions. The sunspot is uniformly distributed over $[0, 1]$ independently of the sequence of action profiles and sunspots realized in the past. Each player observes the realized sunspot without any cost.

The sequence of events in each period is summarized as follows. First, each player i simultaneously chooses an action $a_i \in A_i$ and decides whether to monitor the action chosen by the opponent player or not. Finally, a sunspot is realized.

Player i's history at the beginning of period $t \geq 2$ is a sequence of his own actions, his observations about the other player's actions and realized sunspots up to period $t-1$. Formally, it is a sequence

$$h_i^t = (a_i^s, b_i^s, x^s)_{s=1}^{t-1} \in (A_i \times O_i \times [0, 1])^{t-1}.$$

For $t \geq 1$, let $\mathcal{H}_i^t$ denote the set of all player i's histories at the beginning of period t. Let $\mathcal{H}_i^1$ be an arbitrary singleton set. For any set K, let $\Delta(K)$ be the set of probability distributions over K. A strategy of player i is a function of his private history to distributions over $A_i \times B_i$, $\sigma_i : \cup_{s=1}^{\infty} \mathcal{H}_i^t \to \Delta(A_i \times B_i)$.

A strategy profile $\sigma = (\sigma_1, \sigma_2)$ generates a probability distribution over the set of outcomes, $(A \times B_i)^{\infty}$. Given an outcome $(a^s, b_i^s)_{s=1}^{\infty}$, player i's discounted average payoff is

$$(1 - \delta) \sum_{s=1}^{\infty} \delta^s \tilde{u}_i(a^s, b_i^s),$$

where $\delta \in (0, 1)$ is a discount factor common between player 1 and player 2. Players maximize the expected discounted average payoffs. We use sequential equilibrium as a solution concept.

We assume that each player observes nothing if he does not incur a cost. This implies that each player does not receive the stage game payoffs until the infinitely repeated game "ends", and he receives them after the infinitely repeated game ends. In the basic interpretation of infinitely repeated games, the infinitely repeated game is assumed to be continued permanently. However, if we consider the discount factor as a probability that the infinitely repeated game continues, then the interpretation that each player receives his payoff after the infinitely repeated game ends is less problematic. This assumption is extreme, and it eliminates all the issues about the monitoring structure besides costly observation from our model.

5.3 An efficiency

In this section, we show an efficiency result when public randomization is available.

Proposition 5.1 *Suppose that Assumption 5.1 is satisfied. For any* $\varepsilon > 0$*, there exists* $\overline{\delta} \in (0,1)$ *and* $\overline{\lambda} > 0$ *such that for any* $\delta \in [\overline{\delta}, 1)$ *and any* $\lambda \in (0, \overline{\lambda}]$*, there exists a sequential equilibrium whose payoff vector* (v_1^*, v_2^*) *satisfies that* $|v_i^* - 1| < \varepsilon$ *for* $i = 1, 2$*.*

Proposition 5.1 holds even when we focus on automata strategies as equilibrium strategies.

Definition 5.1 An automaton $(S_i, s_i^\phi, f_i, T_i)$ consists of the set of state S_i, an initial state $s_i^\phi \in S_i$, an output function $f_i : S_i \to \Delta(A_i \times B_i)$, and a transition function $T_i : S_i \times A_i \times B_i \to S_i$. An automata is said to be automation sequential equilibrium if the automaton induces a sequential equilibrium strategy.

In fact, we show Proposition 5.1 by constructing automation sequential equilibrium.

Proof. First, we define automata and a belief system. Next, we show that the belief system is consistent with the strategy induced by the automata. Finally, we show that the pair of the strategy and the belief system is a sequential equilibrium and prove Proposition 5.3

Let us consider the following automaton with states s_i^ϕ, s_i^C and s_i^D.

The output function f_i and the transition function T_i at state s_i^ϕ is defined as follows. Player i chooses (C_i, M_i) with probability $(1-\alpha)(1-\beta)$, and chooses (C_i, N_i) with probability $(1-\alpha)\beta$, and chooses (D_i, N_i) with the remaining probability α. When sunspot satisfies that $x > \underline{x}$, the state remains the same. The state moves to s_i^C if sunspot satisfies that $x \leq \underline{x}$ and player i played C_i and observed C_j. Otherwise, the state moves to state s_i^D. The output function from state s_i^ϕ is summarized as follows.

$$f_i(s_i^\phi) = (1-\alpha)(1-\beta)[(C_i, M_i)] + (1-\alpha)\beta[(C_i, N_i)] + \alpha[(D_i, N_i)].$$

Transition function from state s_i^ϕ is summarized as follows.

$$T_i(s_i^\phi, a_i, o_i, x) = \begin{cases} s_i^C, & \text{if } x \leq \underline{x} \text{ and } (a_i, o_i) = (C_i, C_j), \\ s_i^D, & \text{if } x \leq \underline{x} \text{ and } (a_i, o_i) \neq (C_i, C_j), \\ s_i^\phi, & \text{if } x > \underline{x}, \end{cases}$$

where $(a_i, o_i) \in A_i \times O_i$.

The output function f_i and the transition function T_i at state s_i^C is defined as follows. Player i chooses (C_i, M_i) with probability $1-\beta$, and chooses (C_i, N_i) with the remaining probability β. When sunspot satisfies that $x > \hat{x}$, the state moves to state s_i^ϕ. The state moves to s_i^C if sunspot satisfies that $x \leq \hat{x}$ and player i played C_i and observed C_j. Otherwise, the state moves to state s_i^D. The output function from state s_i^C is summarized as follows.

$$f_i(s_i^C) = (1-\beta)[(C_i, M_i)] + \beta[(C_i, N_i)].$$

Transition function from state s_i^C is summarized as follows.

$$T_i(s_i^C, a_i, o_i, x) = \begin{cases} s_i^C, & \text{if } x \leq \hat{x} \text{ and } (a_i, o_i) = (C_i, C_j), \\ s_i^D, & \text{if } x \leq \hat{x} \text{ and } (a_i, o_i) \neq (C_i, C_j), \\ s_i^\phi, & \text{if } x > \hat{x}, \end{cases}$$

where $(a_i, o_i) \in A_i \times O_i$.

The output function f_i and the transition function T_i at state s_i^D is defined as follows. Player i chooses action (D_i, N_i) with probability 1. The state moves to s_i^C if sunspot satisfies that $x \leq \hat{x}$ and player i played C_i and observed C_j (Note that this event is off the path). Otherwise, the state remains the same. The output function from state s_i^D is summarized as follows.

$$f_i(s_i^D) = 1 \cdot [(D_i, N_i)].$$

Transition function from state s_i^D is summarized as follows.

$$T_i(s_i^D, a_i, o_i, x) = \begin{cases} s_i^C, & \text{if } x \leq \hat{x} \text{ and } (a_i, o_i) = (C_i, C_j), \\ s_i^D, & \text{if } x \leq \hat{x} \text{ and } (a_i, o_i) \neq (C_i, C_j), \\ s_i^\phi, & \text{if } x > \hat{x}, \end{cases}$$

where $(a_i, o_i) \in A_i \times O_i$.

In what follows, we assume that $\alpha \in (0, 1)$ and $\beta \in (0, 1)$ hold and we define a belief system. Let $\psi_i^t(h_i^t)$ be the belief of player i at the beginning of period t over $\mathcal{H}_j^t$ given h_i^t. We call $\psi_i \equiv (\psi_i^t)_{t=1}^\infty$ player i's belief and $\psi = (\psi_1, \psi_2)$ a belief system.

We define a belief system $\hat{\psi} = (\hat{\psi}_1, \hat{\psi}_2)$ which is consistent with the strategy. Belief system $\hat{\psi}$ is generated by the following trembling of the output functions.

$$f_i^\gamma(s_i^\phi) = (1-\gamma)\{(1-\alpha)(1-\beta)[(C_i, M_i)] + (1-\alpha)\beta[(C_i, N_i)] + \alpha[(D_i, N_i)]\} + \gamma[(D_i, M_i)].$$
$$f_i^\gamma(s_i^C) = (1-2\gamma)\{(1-\beta)[(C_i, M_i)] + \beta[(C_i, N_i)]\} + \gamma[(D_i, N_i)] + \gamma[(D_i, M_i)].$$
$$f_i^\gamma(s_i^D) = \gamma\{(1-\beta)[(C_i, M_i)] + \beta[(C_i, N_i)]\} + (1-2\gamma)[(D_i, N_i)] + \gamma[(D_i, M_i)]$$

The output function f_i^γ converges to f_i as γ goes to zero. Four-tuple $(f_1^\gamma, f_2^\gamma, T_1, T_2)$ generates a belief system $\psi_{1,\gamma}$ for any $\gamma > 0$. We define player i's belief:$\hat{\psi}_i \equiv \lim_{\gamma \downarrow 0} \psi_{1,\gamma}$.

Each player i's strategy is generated by the automaton f_i which depends only on player i's state. Hence, it is sufficient to consider beliefs over the set of the other player's states instead of beliefs over the set of the other player's histories in order to examine player i's sequential rationality. By the construction of $\hat{\psi}^t$, we find the following facts about beliefs over the set of states.

First, we consider the set of histories at period t in which player i's state at period t is state s_i^ϕ. We can divide the set into the following three sets: (1) initial history; (2) the set of histories in which the state of player i at period $t-1$ is s_i^ϕ and $x > \underline{x}$ is realized in period $t-1$; (3) the set of histories in which player i's state at period $t-1$ is either s_i^C or s_i^D and x^t is greater than $\hat{x}$.

For any sequence of realized sunspots, player i's state is s_i^ϕ if and only if player j's state is s_j^ϕ by the construction of transition functions. Thus, for any period t, both on and off the path, player i believes that player j's state at period t is s_j^ϕ with probability one.

Second, we consider state s_i^C. We divide the set of histories at period t in which player i's state at period t is s_i^C into three sets: (1) the set of histories in which player i's state at period $t-1$ was s_i^ϕ, player i chose action C_i and observed C_j in period $t-1$ and x^t was not greater than $\underline{x}$; (2) the set of histories in which the state of player i at period $t-1$ is s_i^C, player i observed (C_i, C_j) in period $t-1$ and $x \leq \hat{x}$ was realized in period $t-1$; (3) the set of histories in which player i's state at period $t-1$ was s_i^D, player i chose action C_i and observed C_j in period $t-1$ and x^t was not greater than $\hat{x}$ (off the path).

At both state s_j^ϕ and state s_j^C, player j chooses action C_j with a positive probability, and player j observes player i's action with probability $1-\beta$ given action C_j. In addition, by the construction of tremble, player i believes that when player j chooses action C_j at state s_j^D (off the path), player j observes player i's action with probability $1-\beta$.

Consider a history is in the set of (1). Then, player i in period $t-1$ believed that player j's state at period $t-1$ is s_j^ϕ with probability one. Hence, player i in period t believes that player j's state at period t is s_j^C with probability $1-\beta$ and s_j^D with probability β because player i believes that player j in period $t-1$ observed player i's action C_i with probability $1-\beta$.

Next, suppose that the state of player i in period $t-1$ is either s_i^C or s_i^D. Then, player i in period $t-1$ believed that player j's state at period $t-1$ is s_i^C or s_i^D with probability one. Player i believes that player j's state in period t is s_j^C with probability $1-\beta$ and s_j^D with probability β because given observation C_j, player i believes that player j in period $t-1$ observed player i's action C_i with probability $1-\beta$ at each state.

Finally, we consider the set of histories in which player i's state at period t is s_i^D. The set can be divided into the following sets. The first set is the set of histories in which player i's state is s_i^ϕ in period period $t-1$, either player i chooses action D_i in period $t-1$ or player i observes action D_j in period $t-1$ and $x \leq \underline{x}$ is realized in period $t-1$. The second set is the set of histories in which the state of player i at period $t-1$ is either s_i^C or s_i^D, either player i chooses action D_i in period $t-1$ or player i observes action D_j in period $t-1$ and x^{t-1} is not greater than $\hat{x}$. Let us denote by $t_i^\phi (< t)$ the latest period at which player i's state is s_i^ϕ. Then, the third set is the set of histories in which player i chooses C_i, player i does not observe player j's action in period $\tilde{t} = t_i^\phi, t_i^\phi + 1, \ldots, t-1$, sunspot $x^{t_i^\phi}$ is not greater than $\underline{x}$ and $x^{\tilde{t}} \leq \hat{x}$ holds for $\tilde{t} = t_i^\phi + 1, t_i^\phi + 2, \ldots, t-1$. Let us denote by $t_i^C (< t)$ the latest period at which player i's state is s_i^C. The final set is the set of histories in which player i chooses C_i, player i does not observe player j's action in period $\tilde{t} = t_i^C, t_i^C + 1, \ldots, t-1$ and $x \leq \hat{x}$ is realized in period $\tilde{t} = t_i^C, t_i^C + 1, \ldots, t-1$.

For any history in the set of (1)–(2), player i believes that player j's state is s_j^D with probability one. We consider histories in the set of (3). Suppose that $t_i^\phi = t-1$ holds. Then, player i in period $t-1$ believed that player j's state in period $t-1$ was s_j^ϕ. Therefore, player i in period t believes that player j's state at period t is s_j^C with probability $(1-\alpha)(1-\beta)$ and s_j^D with probability $1-(1-\alpha)(1-\beta)$ because player i chose C_i and player i did not observe player j's action in period $t-1$. Next, suppose that $t - t_i^\phi > 1$. Then, in the same way as the case for $t_i^\phi = t-1$, player i

in period t_i^ϕ believed that player j's state at period t_i^ϕ was s_j^ϕ. Thus, player j's state in period t is s_j^C if and only if player j chose C_j and player j observed player i's action in period $\tilde{t} = t_i^\phi, t_i^\phi + 1, \ldots, t-1$. Hence, player i in period t believes that player j's state at period t is s_j^C with probability $(1-\alpha)(1-\beta)^{t-t_i^\phi}$ and s_j^D with probability $1-(1-\alpha)(1-\beta)^{t-t_i^\phi}$.

We can consider histories in the set of (4) in the same way as histories in the set of (3). First, suppose that $t_i^C = t-1$ holds. Then, player i's state at period $t-1$ was s_i^C and player i in period $t-1$ believed that player j's state at period $t-1$ was s_j^C with probability $1-\beta$ and s_j^D with probability β. On the path, player j moves to state s_j^C in period t if and only if player j observed player i's action at state s_i^C in period $t-1$. Thus, player i in period t believes that player j's state at period t is s_j^C with probability $(1-\beta)^2$ and s_j^D with probability $1-(1-\beta)^2$. Next, suppose that $t-t_i^C > 1$. Then, in the same way as the case for $t_i^C = t-1$, player i's state in period t_i^C was s_i^C and player i in period t_i^C believed that player j's state at period $t-1$ was s_j^C with probability $1-\beta$ and s_j^D with probability β. Player j moves to state s_j^C in period t if and only if player j observed player i's action at state s_i^C in period $\tilde{t} = t_i^C, t_i^C+1, \ldots, t-1$. This event happens with probability $(1-\beta)^{t-t_i^C}$ when player j's state at period t_i^C is s_j^C. Therefore, player i in period t believes that player j's state in period t is s_j^C with probability $(1-\beta)^{t-t_i^C+1}$ and s_j^D with probability $1-(1-\beta)^{t-t_i^C+1}$.

Next, we show that it is optimal for each player to follow the automata given the belief system. Let V_i^C be the continuation payoff when player i believes that the state of player j is s_j^C with probability $1-\beta$, s_j^D with probability β. Let us denote by V_i^D the continuation payoff when player i believes that the state of player j is s_j^D. Let V_i^ϕ be the payoff given this automata. If the automata is a sequential equilibrium, then, the continuation payoff V_i^C and V_i^D are given by

$$V_i^C = (1-\delta)[(1-\beta) - \beta\ell - \lambda] + \delta\hat{x}(1-\beta)V_i^C + \delta\hat{x}\beta V_i^D + \delta(1-\hat{x})V_i^\phi, \tag{5.1}$$

$$V_i^D = \delta\hat{x}V_i^D + \delta(1-\hat{x})V_i^\phi. \tag{5.2}$$

First, we consider player i's best response at state s_i^ϕ. If player i chooses C_i and observes C_j and sunspot satisfies $x^1 \leq \underline{x}$ in period 1, then player i in period 2 believes that the state of player j is s_j^C with probability $1-\beta$, s_j^D with probability β. That is, the continuation payoff from period 2 is V_i^C. If player i chooses D_i and sunspot satisfies $x^1 \leq \underline{x}$ in period 1, then player i believes that the state of player j is s_j^D with probability one. That is, the continuation payoff from period 2 is V_i^D.

In addition, (D_i, M_i) is never a best response at any history because the continuation strategy when he chooses action D_i and observes C_j and the continuation strategy when he chooses action D_i and observes D_j lead to the same result. Hence, it is optimal for player i in period 1 to follow the automaton strategy if the following is held.

$$
\begin{aligned}
&(1-\delta)[(1-\alpha)-\alpha\ell-\lambda]+\delta\underline{x}(1-\alpha)V_i^C+\delta\underline{x}\alpha V_i^D+\delta(1-\underline{x})V_i^\phi \\
=&(1-\delta)[(1-\alpha)-\alpha\ell]+\delta\underline{x}(1-\alpha)(1-\beta)(1+g) \qquad (5.3)\\
&+\delta^2\underline{x}\hat{x}V_i^D+\delta(1-\underline{x})V_i^\phi+\delta^2\underline{x}(1-\hat{x})V_i^\phi \\
=&(1-\delta)(1-\alpha)(1+g)+\delta\underline{x}V_i^D+\delta(1-\underline{x})V_i^\phi \qquad (5.4)
\end{aligned}
$$

Next, we consider player i's best response at state s_i^C. If player i's state is s_i^C, then player i believes that player j's state is s_j^C with probability $1-\beta$, and s_j^D with probability β. Considering that (D_i, M_i) is never a best response for any history, we find that it is optimal for player i to follow the automaton strategy at state s_i^C, if the following is held.

$$
\begin{aligned}
&(1-\delta)[(1-\beta)-\beta\ell-\lambda]+\delta\hat{x}(1-\beta)V_i^C+\delta\hat{x}\beta V_i^D+\delta(1-\hat{x})V_i^\phi \\
=&(1-\delta)[(1-\beta)-\beta\ell]+\delta\hat{x}(1-\beta)^2(1+g) \qquad (5.5)\\
&+\delta^2\hat{x}^2V_i^D+\delta(1-\hat{x})V_i^\phi+\delta^2\hat{x}(1-\hat{x})V_i^\phi \\
\geq&(1-\delta)[(1-\beta)(1+g)]+\delta\hat{x}V_i^D+\delta(1-\hat{x})V_i^\phi. \qquad (5.6)
\end{aligned}
$$

We rewrite the inequalities (5.3)–(5.6). Substituting (5.1) and (5.2) into righthand side of (5.5) repeatedly, we obtain

$$
\begin{aligned}
&(1-\delta)[(1-\beta)-\beta\ell-\lambda]+\delta\hat{x}(1-\beta)V_i^C+\delta\hat{x}\beta V_i^D+\delta(1-\hat{x})V_i^\phi \\
=&(1+\delta\hat{x}\beta)(1-\delta)[(1-\beta)-\beta\ell-\lambda]+(\delta\hat{x})^2(1-\beta)V_i^C+(\delta\hat{x})^2\beta V_i^D+(1+\delta\hat{x})\delta(1-\hat{x})V_i^\phi \\
=&\frac{[(1-\beta)-\beta\ell-\lambda]}{1-\delta\hat{x}(1-\beta)}+\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}V_i^\phi
\end{aligned}
$$

In the same way, inequalities (5.5)–(5.6) can be rewritten as follows.

$$
\begin{aligned}
&\frac{(1-\delta)[(1-\beta)-\beta\ell-\lambda]}{1-\delta\hat{x}(1-\beta)}+\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}V_i^\phi \\
=&(1-\delta)[(1-\beta)-\beta\ell]+(1-\delta)\delta\hat{x}(1-\beta)^2(1+g)+\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}V_i^\phi \\
\geq&(1-\delta)(1-\beta)(1+g)+\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}V_i^\phi
\end{aligned}
$$

Let us define a discount factor $\hat{\delta}$ and a continuation payoff w as follows.

$$
\begin{aligned}
\hat{\delta}\equiv&\delta\hat{x} \\
w\equiv&(1-\hat{\delta})[(1-\beta)-\beta\ell-\lambda]+\hat{\delta}(1-\beta)w=\frac{(1-\hat{\delta})[(1-\beta)-\beta\ell-\lambda]}{1-\hat{\delta}(1-\beta)}
\end{aligned}
$$

Then, inequalities (5.5)– (5.6) are equivalent to the following.

$$
w=(1-\hat{\delta})[(1-\beta)-\beta\ell]+(1-\hat{\delta})\hat{\delta}(1-\beta)^2(1+g)\geq(1-\hat{\delta})(1-\beta)(1+g).
$$

In the same way, equalities (5.3)– (5.4) can be rewritten as follows.

$$
\begin{aligned}
V_i^\phi=&(1-\delta)[(1-\alpha)-\alpha\ell-\lambda]+\delta\underline{x}(1-\alpha)w+\left(\delta(1-\underline{x})+\delta\underline{x}\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}\right)V_i^\phi \qquad (5.7)\\
=&(1-\delta)[(1-\alpha)-\alpha\ell]+(1-\delta)\delta\underline{x}(1-\alpha)\beta(1+g)+\left(\delta(1-\underline{x})+\delta\underline{x}\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}\right)V_i^\phi \\
=&(1-\delta)(1-\alpha)(1+g)+\left(\delta(1-\underline{x})+\delta\underline{x}\frac{\delta(1-\hat{x})}{1-\delta\hat{x}}\right)V_i^\phi.
\end{aligned}
$$

Let us define a discount factor $\underline{\delta}\equiv\delta\overline{x}$. Then, equations (5.3)– (5.4) are equivalent to the following.

$$
(1-\hat{\delta})[\alpha-(1-\alpha)\ell-\lambda]+\underline{\delta}\alpha w=(1-\hat{\delta})[\alpha-(1-\alpha)\ell]+\underline{\delta}\alpha\beta(1+g)=(1-\hat{\delta})\alpha(1+g)
$$

Therefore, we have the following fact.

Fact 5.1 *Fix a discount factor* $\delta \in (0,1)$ *and monitoring cost* $\lambda > 0$. *Fix a 5-tuple* $(\alpha, \beta, \underline{\delta}, \hat{\delta}, w)$. *From equality* (5.7)*, the payoff* V_i^ϕ *is given by*

$$V_i^\phi = \frac{1}{1-\underline{\delta}+\hat{\delta}}\left[(1-\hat{\delta})[(1-\alpha)-\alpha\ell-\lambda]+\underline{\delta}(1-\alpha)w\right].$$

It is optimal for player i *to follow the automaton strategy at state* s_i^ϕ *if and only if it holds that*

$$\begin{aligned}&(1-\hat{\delta})[(1-\alpha)-\alpha\ell-\lambda]+\underline{\delta}(1-\alpha)w && (5.8)\\ =&(1-\hat{\delta})[(1-\alpha)-\alpha\ell]+(1-\hat{\delta})\underline{\delta}(1-\alpha)(1-\beta)(1+g) && (5.9)\\ =&(1-\hat{\delta})(1-\alpha)(1+g). && (5.10)\end{aligned}$$

It is optimal for player i *to follow the automaton strategy at state* s_i^C *if and only if it holds that*

$$\begin{aligned}w =&(1-\hat{\delta})[(1-\beta)-\beta\ell-\lambda]+\hat{\delta}(1-\beta)w && (5.11)\\ =&(1-\hat{\delta})[(1-\beta)-\beta\ell]+(1-\hat{\delta})\hat{\delta}(1-\beta)^2(1+g) && (5.12)\\ \geq&(1-\hat{\delta})(1-\beta)(1+g). && (5.13)\end{aligned}$$

Fix $\varepsilon > 0$. In what follows, we show that if we fix any sufficiently small $\tilde{\varepsilon}(<\varepsilon)$, then there exists $\overline{\delta} \in (0,1)$ and $\overline{\lambda} > 0$ such that for any $\delta \in [\overline{\delta}, 1)$ and any $\lambda \in (0, \overline{\lambda}]$, there exists a sequential equilibrium whose payoff vector (v_1^*, v_2^*) satisfies that $|v_i^* - 1| < \tilde{\varepsilon}$ for $i = 1, 2$. This statement is one of the sufficient conditions that Proposition 5.1 holds.

Fix any sufficiently small $\tilde{\varepsilon}(<\varepsilon)$. We define $\overline{\delta}$ and $\overline{\lambda}$:

$$\begin{aligned}\overline{\delta} \equiv&\frac{g+2\tilde{\varepsilon}}{1+g},\\ \overline{\lambda} \equiv&\frac{\ell(g-\ell)}{18(1+g)^3}\tilde{\varepsilon}.\end{aligned}$$

Fix any $\delta \in [\overline{\delta}, 1)$ and $\lambda \in (0, \overline{\lambda})$. First, we fix $w = 1 - \tilde{\varepsilon}$. We define $(\alpha, \beta, \underline{\delta}, \hat{\delta})$ as a solution of simultaneous equations (5.9)–(5.12). We have the followings.

$$\begin{aligned}\beta =&\frac{(1-\hat{\delta})(\tilde{\varepsilon}-\lambda)}{(1-\hat{\delta})(1+\ell)+\hat{\delta}(1-\tilde{\varepsilon})},\\ \alpha =&\frac{g-\hat{\delta}(1-\beta)^2(1+g)}{g-\ell},\\ \underline{\delta} =&\frac{1-\beta}{1-\alpha}\hat{\delta}.\end{aligned} \tag{5.14}$$

All we have to do is to show the followings. (1) The strategy is well-defined. That is, $0 < \alpha, \beta, \underline{x}, \hat{x} < 1$ holds, and (2) player i's best response at state s_i^C is (C_i, M_i) and (C_i, N_i), that is, inequality (5.13) holds, and (3) player i's unique best response at state s_i^D is (D_i, N_i), (4) the equilibrium payoff V_i^{ϕ} is greater than $1-\tilde{\varepsilon}$.

To show that the strategy is well-defined, we prove $\alpha \in (0,1)$, $\beta \in (0,1)$, $\hat{\delta} \in (0,1)$ and $\underline{\delta} \in (0,1)$. First, we prove $\beta \in (0,1)$. We obtain $\beta > 0$ if $\hat{\delta} \in (0,1)$ holds because $\lambda < \overline{\lambda} < \tilde{\varepsilon}$ hold and $\tilde{\varepsilon}$ is small. We also derive $\beta < \tilde{\varepsilon}$ for small $\tilde{\varepsilon}$ if $\hat{\delta} \in (0,1)$ holds as follows.

$$\tilde{\varepsilon}-\beta = \frac{\tilde{\varepsilon}\{(1-\hat{\delta})\ell+\hat{\delta}(1-\tilde{\varepsilon})\}+(1-\hat{\delta})\lambda}{(1-\hat{\delta})(1+\ell)+\hat{\delta}(1-\tilde{\varepsilon})} > 0.$$

Hence, we obtain $\beta \in (0, \tilde{\varepsilon}) \subset (0,1)$ because $\tilde{\varepsilon}$ is sufficently small.

Second, we show $\alpha \in (0,1)$. From equality (5.12), we have

$$(1-\hat{\delta})(1-\beta)(1+g) = w - \frac{1-\hat{\delta}}{\hat{\delta}(1-\beta)}\lambda. \tag{5.15}$$

Using equality (5.12), we have

$$\begin{aligned}&(1-\hat{\delta})(1-\alpha)(1+g)\\ =&(1-\hat{\delta})(1-\beta)(1+g)+(1-\hat{\delta})(\beta-\alpha)(1+g)\\ =&w-\frac{1-\hat{\delta}}{\hat{\delta}(1-\beta)}\lambda+(1-\hat{\delta})(\beta-\alpha)(1+g).\end{aligned}$$

Using equality (5.12), we have

$$\begin{aligned}
&(1-\hat{\delta})[(1-\alpha)-\alpha\ell]+(1-\hat{\delta})\underline{\delta}(1-\alpha)(1-\beta)(1+g)\\
=&(1-\hat{\delta})[(1-\beta)-\beta\ell]+(1-\hat{\delta})\hat{\delta}(1-\beta)^2(1+g)+(1-\hat{\delta})(\beta-\alpha)(1+\ell)\\
=&w+(1-\hat{\delta})(\beta-\alpha)(1+\ell).
\end{aligned}$$

From the above two equalities, we have.

$$\beta-\alpha=\frac{\lambda}{\hat{\delta}(1+\ell)(g-\ell)}>0. \tag{5.16}$$

Assumption 5.1–(iii) ensures strictly inequality. We obtain $\alpha(<\beta<\tilde{\varepsilon})<1$.

To prove that α is positive, we show the following lemma.

Lemma 5.1 *Suppose that Assumption 5.1 is satisfied. Then, the followings hold.*

$$\hat{\delta}\in\left(\frac{g}{1+g},\frac{g+2\tilde{\varepsilon}}{1+g}\right),\quad \textit{and} \tag{5.17}$$

$$\beta>\frac{\tilde{\varepsilon}-\lambda}{2(1+g)}. \tag{5.18}$$

Proof. By equality (5.12), we have

$$\frac{1-\hat{\delta}}{\hat{\delta}}\lambda-(1-\beta)\{1-\tilde{\varepsilon}-(1-\hat{\delta})(1+g)\}=0. \tag{5.19}$$

The lefthand side of (5.19) is positive if $\hat{\delta}=\frac{g}{1+g}$. If $\hat{\delta}=\frac{g+2\tilde{\varepsilon}}{1+g}$, then the lefthand side of (5.19) is negative because $\lambda(\le\overline{\lambda})<\tilde{\varepsilon}$ and $\beta<\tilde{\varepsilon}$ and $\tilde{\varepsilon}$ is sufficiently small. Therefore, there exists $\hat{\delta}\in\left(\frac{g}{1+g},\frac{g+2\tilde{\varepsilon}}{1+g}\right)$ that satisfies equality (5.12).

Then, it holds that

$$\begin{aligned}
\beta=&\frac{(1-\hat{\delta})(\tilde{\varepsilon}-\lambda)}{(1-\hat{\delta})(1+\ell)+\hat{\delta}(1-\tilde{\varepsilon})}\\
>&\frac{\tilde{\varepsilon}-\lambda}{(1+\ell)+g(1-\tilde{\varepsilon})}\\
>&\frac{\tilde{\varepsilon}-\lambda}{2(1+g)}.
\end{aligned}$$

□

From equation (5.16) and the above lemma, we have

$$
\begin{aligned}
\alpha =& \beta - \frac{\lambda}{\hat{\delta}(1-\beta)(g-\ell)} \\
>& \beta - \frac{\lambda}{\hat{\delta}(g-\ell)} \\
>& \frac{\tilde{\varepsilon}-\lambda}{2(1+g)} - \frac{1+g}{g(g-\ell)}\lambda \\
>& 0.
\end{aligned}
\tag{5.20}
$$

The last inequality follows from $\lambda \leq \overline{\lambda}$. Thus, we have $\alpha \in (0, \beta) \subset (0, 1)$.

Third, we show $\hat{x} \in [0, 1]$. By the above lemma, we have $\hat{\delta} \in \left(\frac{g}{1+g}, \frac{g+2\tilde{\varepsilon}}{1+g}\right)$. That is, $0 < \hat{\delta}(< \overline{\delta}) < \delta$. It means that $\hat{x} \in (0, 1)$.

Finally, we show $\underline{x} \in [0, 1]$. By inequality (5.16) and $0 < \alpha < \beta < 1$, we have $0 < \underline{\delta} < \hat{\delta} < \delta$. It implies that that $\hat{x} \in (0, 1)$.

Equality (5.15) implies that strict inequality holds in inequality (5.13). That is, it is proved that it is optimal for each player to follow the automata at state s_i^C.

Next, we prove that it is optimal for each player to follow the automata at state s_i^D. To prove it, we consider the following situation to show that (D_i, N_i) is a best response when the automaton prescribes to move to state s_i^D. Suppose that a private history h_i^2 at which $(a_i^1, o_i^1) = (C_i, N_i)$ were realized in period 1. Assume that player i can observe player j costless in period $t \geq 2$.

In such a situation, player i's optimal strategy from period 2 is the grim trigger strategy in which he chooses (C_i, M_i) if (C_i, C_j) is realized in period $1, 2, \ldots,$ he chooses (D_i, M_i) otherwise. The continuation payoff from period 2 is given by

$$
\hat{w} = w + \frac{1-\hat{\delta}}{1-\hat{\delta}(1-\beta)}\lambda.
$$

The above payoff is greater than the payoff when he chooses (C_i, N_i). Consider a private history at which the automata prescribes to move to state s_i^D. Given such private history, player i has the following belief over S_j. The state of player j is s_j^ϕ with probability zero, it is s_j^C with probability $p(\leq (1-\alpha)(1-\beta))$ and it is s_j^D with probability $1-p$. Hence, if he chooses (C_i, N_i), then he has

$$\begin{aligned}&(1-\hat{\delta})[p-(1-p)\ell]+(1-\hat{\delta})\hat{\delta}p(1-\beta)(1+g)\\<&(1-\hat{\delta})[(1-\beta)-\beta\ell]+(1-\hat{\delta})\hat{\delta}(1-\beta)^2(1+g)=w<\hat{w}.\end{aligned}$$

Thus, (C_i, M_i) is more profitable for player i than (C_i, N_i). Therefore, we compare the payoff when player i chooses (D_i, N_i) with the payoff when player i chooses (C_i, M_i). Choosing (D_i, N_i) is more profitable than choosing (C_i, M_i) for player i when the following is positive.

$$(1-\hat{\delta})p(1+g)-\left[(1-\hat{\delta})\left\{p-\ell(1-p)\right\}+\hat{\delta}p\hat{w}\right] \tag{5.21}$$

The above is a decreasing function of p because the following holds.

$$\hat{\delta}\hat{w}-(1-\hat{\delta})(g-\ell)>\frac{1}{1+g}\left\{g(1-\tilde{\varepsilon})-(g-\ell)\right\}>\frac{1}{1+g}(\ell-\tilde{\varepsilon})>0.$$

We show that (5.21) is positive when $p=(1-\alpha)(1-\beta)$.

$$\begin{aligned}&(1-\hat{\delta})(1-\alpha)(1-\beta)(1+g)\\&-\left[(1-\hat{\delta})\left\{(1-\alpha)(1-\beta)-\ell(\alpha+\beta-\alpha\beta)\right\}+\hat{\delta}(1-\alpha)(1-\beta)\hat{w}\right]\\=&\alpha[(1-\hat{\delta})(1-\beta)(\ell+\tilde{\varepsilon})-\beta w]-\frac{1-\hat{\delta}}{\hat{\delta}(1-\beta)}\lambda-\left\{1-\hat{\delta}\alpha(1-\beta)\right\}\frac{1-\hat{\delta}}{1-\hat{\delta}(1-\beta)}\lambda\\&\text{(by inequality (5.20))}\\>&\left\{\frac{\tilde{\varepsilon}-\lambda}{2(1+g)}-\frac{1+g}{g(g-\ell)}\lambda\right\}[(1-\hat{\delta})(1-\beta)\ell-\beta]-\frac{2(1-\hat{\delta})}{\hat{\delta}}\lambda-\lambda\\&\text{(by }\hat{\delta}\in(\underline{\delta},\overline{\delta})\text{ and }\beta<\tilde{\varepsilon})\\>&\left\{\frac{\tilde{\varepsilon}-\lambda}{2(1+g)}-\frac{1+g}{g(g-\ell)}\lambda\right\}\left\{\frac{\ell}{1+g}-\left(1+\frac{2\tilde{\varepsilon}(1-\tilde{\varepsilon})\ell}{1+g}\right)\tilde{\varepsilon}\right\}-2\frac{1+g}{g}\lambda\end{aligned}$$

The values of $\tilde{\varepsilon}(\tilde{\varepsilon}-\lambda)$, $\tilde{\varepsilon}\lambda$ and $\tilde{\varepsilon}^2$ are relatively smaller than $\tilde{\varepsilon}$ when $\tilde{\varepsilon}$ is close enough to zero: $\lim_{\tilde{\varepsilon}\to 0}\frac{\tilde{\varepsilon}(\tilde{\varepsilon}-\lambda)}{\tilde{\varepsilon}}\leq\lim_{\tilde{\varepsilon}\to 0}\frac{\tilde{\varepsilon}\lambda}{\tilde{\varepsilon}}\leq\lim_{\tilde{\varepsilon}\to 0}\frac{\tilde{\varepsilon}^2}{\tilde{\varepsilon}}=0$. Therefore, if $\tilde{\varepsilon}$ is sufficiently small, then we have

$$
\begin{aligned}
&\left\{\frac{\tilde{\varepsilon}-\lambda}{2(1+g)}-\frac{1+g}{g(g-\ell)}\lambda\right\}\left\{\frac{\ell}{1+g}-\left(1+\frac{2\tilde{\varepsilon}(1-\tilde{\varepsilon})\ell}{1+g}\right)\tilde{\varepsilon}\right\}-2\frac{1+g}{g}\lambda \\
>&\left[\left\{\frac{\tilde{\varepsilon}-\lambda}{2(1+g)}-\frac{1+g}{g(g-\ell)}\lambda\right\}\frac{\ell}{1+g}-2\frac{1+g}{g}\lambda\right]-\frac{1+g}{g}\lambda \\
=&\frac{\ell}{2(1+g)^2}\tilde{\varepsilon}-\frac{\ell}{2(1+g)^2}\lambda-\frac{\ell}{g}\frac{1}{g-\ell}\lambda-\frac{3(1+g)}{g}\lambda \\
=&\frac{\ell}{2(1+g)^2}\tilde{\varepsilon}-\frac{\ell}{2(1+g)^2}\lambda-\frac{\ell}{g}\frac{1}{g-\ell}\lambda-\frac{3(1+g)}{g}\lambda \\
>&\frac{\ell}{2(1+g)^2}\tilde{\varepsilon}-3\left(\frac{3(1+g)}{g-\ell}\right)\lambda>0.
\end{aligned}
$$

The last inequality follows from $\lambda \leq \overline{\lambda}$. Hence, it has been proved that (D_i, N_i) is a best response at state s_i^D.

Finally, we show that the equilibrium payoff is greater than ε. Considering $0 < \alpha < \beta < 1$ and $0 < \underline{\delta} < \hat{\delta} < 1$, we have the equilibrium payoff V_i^{ϕ} as follows.

$$
\begin{aligned}
V_i^{\phi} =&\frac{1}{1-\underline{\delta}+\hat{\delta}}\left[(1-\hat{\delta})[(1-\alpha)-\alpha\ell-\lambda]+\underline{\delta}(1-\alpha)w\right] \\
>&(1-\hat{\delta})[(1-\alpha)-\alpha\ell-\lambda]+\underline{\delta}(1-\alpha)w \\
=&(1-\hat{\delta})[(1-\alpha)-\alpha\ell-\lambda]+\hat{\delta}(1-\beta)w \quad \text{(by equality (5.14))} \\
>&(1-\hat{\delta})[(1-\beta)-\beta\ell-\lambda]+\hat{\delta}(1-\beta)w \\
=&w=1-\tilde{\varepsilon}.
\end{aligned}
$$

Proposition 5.1 has been proved. □

5.4 Conclusion

In this chapter, we analyze an infinitely repeated game with costly observation. Concretely, we analyze a model in which each player can observe the other player's action without noise if he incurs a cost. Otherwise, he observes nothing. Ben-Porath and Kahneman (2003) and Miyagawa et al. (2003) analyze these class of infinitely repeated games and show a sufficient condition for folk theorems. That is, they show a sufficient conditions for efficiency results. Ben-Porath and Kahneman (2003) show that folk theorem holds if communication is available. Miyagawa et al. (2003) show that a sufficient condition that a folk theorem holds when the monitoring cost is sufficiently small. Infinitely repeated prisoner's dilemma is one of the most interesting games. However, the above two studies do not show a sufficient condition for an efficiency result in an infinitely repeated prisoner's dilemma when communication is not available. In this chapter, we show that an efficiency result in an infinitely repeated prisoner's dilemma by constructing a sequential equilibrium when public randomization device is available and the monitoring cost is sufficiently small.

Bibliography

Abreu, D., D. Pearce, and E. Stacchetti (1990) "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, Vol. 58, No. 5, pp. 1041–1063.

Awaya, Y. (2014) "Community Enforcement with Observation Costs," *Journal of Economic Theory*, Vol. 154, pp. 173–186.

Ben-Porath, E. and M. Kahneman (1996) "Communication in Repeated Games with Private Monitoring," *Journal of Economic Theory*, Vol. 70, No. 2, pp. 281–297.

Ben-Porath, E. and M. Kahneman (2003) "Communication in repeated games with costly monitoring," *Games and Economic Behavior*, Vol. 44, No. 2, pp. 227–250.

Bhaskar, V. and E. van Damme (2002) "Moral Hazard and Private Monitoring," *Journal of Economic Theory*, Vol. 102, No. 1, pp. 16–39.

Cole, H. L. and N. R. Kocherlakota (2001) "Dynamic Games with Hidden Actions and Hidden States," *Journal of Economic Theory*, Vol. 98, No. 1, pp. 114–126.

Compte, O. (1998) "Communication in Repeated Games with Imperfect Private Monitoring," *Econometrica*, Vol. 66, No. 3, pp. 587–626.

Cripps, M. W., G. J. Mailath, and L. Samuelson (2004) "Imperfect Monitoring and Impermanent Reputations," *Econometrica*, Vol. 72, No. 2, pp. 407–432.

Cripps, M. W., G. J. Mailath, and L. Samuelson (2007) "Disappearing Private Reputations in Long-run Relationships," *Journal of Economic Theory*, Vol. 134, No. 1, pp. 287–316.

Ekmekci, M. (2011) "Sustainable Reputations with Rating Systems," *Journal of Economic Theory*, Vol. 146, No. 2, pp. 479–503.

Ely, J. C., D. Fudenberg, and D. K. Levine (2008) "When is Reputation Bad?" *Games and Economic Behavior*, Vol. 63, No. 2, pp. 498–526.

Ely, J. C. and J. Välimäki (2002) "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, Vol. 102, No. 1, pp. 84–105.

Ely, J. C. and J. J. Välimäki (2003) "Bad Reputation," *The Quarterly Journal of Economics*, Vol. 118, No. 3, pp. 785–814.

Flesch, J. and A. Perea (2009) "Repeated Games with Voluntary Information Purchase," *Games and Economic Behavior*, Vol. 66, No. 1, pp. 126–145.

Fudenberg, D., D. K. Levine, and E. Maskin (1994) "The Folk Theorem with Imperfect Public Information," *Econometrica*, Vol. 62, No. 5, pp. 997–1039.

Fudenberg, D. and D. K. Levine (1989) "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, Vol. 57, No. 4, pp. 759–778.

Fudenberg, D. and D. K. Levine (1992) "Maintaining a Reputation when Strategies are Imperfectly Observed," *The Review of Economic Studies*, Vol. 59, No. 3, pp. 561–579.

Fudenberg, D. and D. K. Levine (1994) "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, Vol. 62, No. 1, pp. 103–135.

Fudenberg, D. and E. Maskin (1986) "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, Vol. 54, pp. 533–556.

Fudenberg, D. and Y. Yamamoto (2011) "Learning from Private Information in Noisy Repeated Games," *Journal of Economic Theory*, Vol. 146, No. 5, pp. 1733–1769.

Gossner, O. (2011) "Simple Bounds on the Value of a Reputation," *Econometrica*, Vol. 79, No. 5, pp. 1627–1641.

Hörner, J. and W. Olszewski (2006) "The Folk Theorem for Games with Private Almost-Perfect Monitoring," *Econometrica*, Vol. 74, No. 6, pp. 1499–1544.

Kandori, M. and H. Matsushima (1998) "Private Observation, Communication and Collusion," *Econometrica*, Vol. 66, No. 3, pp. 627–652.

Kandori, M. and I. Obara (2004) "Endogenous Monitoring," *Unpublished manuscript, University of Tokyo and UCLA.*

Kaya, A. (2009) "Repeated Signaling Games," *Games and Economic Behavior*, Vol. 66, No. 2, pp. 841–854.

Kreps, D. M. and R. Wilson (1982) "Reputation and Imperfect Information," *Journal of Economic Theory*, Vol. 27, No. 2, pp. 253–279.

Lehrer, E. (1989) "Lower Equilibrium Payoffs in Two-Player Repeated Games with Non-Observable Actions," *International Journal of Game Theory*, Vol. 18, No. 1, pp. 57–89.

Lehrer, E. (1992a) "Correlated Equilibria in Two-Player Repeated Games with Nonobservable Actions," *Mathematics of Operations Research*, Vol. 17, No. 1, pp. 175–199.

Lehrer, E. (1992b) "Two-Player Repeated Games with Nonobservable Actions and Observable Payoffs," *Mathematics of Operations Research*, Vol. 17, No. 1, pp. 200–224.

Liu, Q. (2011) "Information Acquisition and Reputation Dynamics," *The Review of Economic Studies*, Vol. 78, No. 4, pp. 1400–1425.

Liu, Q. and A. Skrzypacz (2014) “Limited Records and Reputation Bubbles,” *Journal of Economic Theory*, Vol. 151, pp. 2–29.

Mailath, G. J. and L. Samuelson (2001) “Who Wants a Good Reputation?” *The Review of Economic Studies*, Vol. 68, No. 2, pp. 415–441.

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

Milgrom, P. and J. Roberts (1982) “Predation, Reputation, and Entry Deterrence,” *Journal of Economic Theory*, Vol. 27, No. 2, pp. 280–312.

Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2003) “Repeated Games with Observation Costs,” *Department of Economics Discussion Papers*, No. 0203-14.

Miyagawa, E., Y. Miyahara, and T. Sekiguchi (2008) “The folk theorem for repeated games with observation costs,” *Journal of Economic Theory*, Vol. 139, No. 1, pp. 192–221.

Monte, D. (2013) “Bounded Memory and Permanent Reputations,” *Journal of Mathematical Economics*, Vol. 49, No. 5, pp. 345–354.

Obara, I. (2009) “Folk Theorem with Communication,” *Journal of Economic Theory*, Vol. 144, No. 1, pp. 120–134.

Sekiguchi, T. (1997) “Efficiency in Repeated Prisoner’s Dilemma with Private Monitoring,” *Journal of Economic Theory*, Vol. 76, No. 2, pp. 345–361.

Selten, R. (1978) “The Chain Store Paradox,” *Theory and Decision*, Vol. 9, No. 2, pp. 127–159.

Sugaya, T. and A. Wolitzky (2014) “Perfect Versus Imperfect Monitoring in Repeated Games,” *working paper, Stanford University.*

Tadelis, S. (1999) “What’s in a Name ? Reputation as a Tradeable Asset,” *The American Economic Review*, Vol. 89, No. 3, pp. 548–563.

Takahashi, S. (2010) “Community Enforcement When Players Observe Partners’ Past Play,” *Journal of Economic Theory*, Vol. 145, No. 1, pp. 42–62.

Wiseman, T. (2005) “A Partial Folk Theorem for Games with Unknown Payoff Distributions,” *Econometrica*, Vol. 73, No. 2, pp. 629–645.

Yamamoto, Y. (2013) “Individual Learning and Cooperation in Noisy Repeated Games,” *The Review of Economic Studies*, Vol. 81, No. 1, pp. 473–500.