博士論文

Cooperation in Various Classes of Repeated Games
(種々の繰り返しゲームにおける協力)

諸岡　千裕

(Chihiro Morooka)

Table of Contents

Acknowledgements

I am deeply indebted to my advisor, Professor Akihiko Matsui, for what I have studied in my master's and doctoral courses. He has spent much time in patiently instructing me in game theory and giving me invaluable advice. My knowledge in this field was very much improved by his support. The model in chapter 1 of this paper is based on his idea. Without his kind encouragement, I could never have completed this work. I am also grateful to my peers, especially Syota Ichihashi, for helpful comments and discussion on the presentation of my research. I would like to thank Editage (www.editage.jp) for English language editing. This work was supported by JSPS KAKENHI Grant Number 17H00978.

General Introduction

Repeated games have been broadly applied to economic models to show that inefficient outcomes can be avoided through repetition. Studies on repeated games usually impose the following assumptions: (a) The number of repetitions is infinite. (b) Each player's opponents are fixed and are not replaced by others throughout the game. (c) Players choose their actions simultaneously. (d) Each player can perfectly observe opponents' previous actions. Although several studies have examined repeated games without these assumptions[1], many questions remain unsolved. Therefore, we consider repeated games where some of these assumptions are not satisfied, and investigate whether and how players can cooperate in such games. This paper consists of four chapters, each of which studies an independent model of repeated games. Our model in each chapter can be applied to explain economic phenomena in the real world.

In chapter 1, we examine two-person overlapping generations (OLG) games. The model in these games does not satisfy (b) in the above assumptions. In OLG games, players in each generation take part in the repeated game for a sufficiently long time, and then are gradually replaced by their successors in the next generation. In the versions of folk theorems in the OLG games proved by Kandori [16], and in subsequent studies, it is required that any (per-period) average equilibrium payoffs are included in the feasible and individually rational one-shot payoff set V. In contrast, in our model of two-person OLG games, it is shown that there exist subgame perfect equilibria where some players can attain possibly high average payoffs outside V. We also guarantee that players who give high one-shot payoffs to other players do not have to decrease their own payoffs, on average, when their lifespan is sufficiently long. Consequently, our result establishes the new folk theorem, which differs from those in prior studies.

For this result, the essential logic we must consider is how to punish the player who failed to reward the other player, who should have obtained high payoffs outside V. In OLG games, a player who received low payoffs outside V earlier in his life in order to contribute to another player can be rewarded before his retirement by the other younger player. In our model of two-person OLG games, players in each generation are classified into three types, each with different roles: a contributed player, a contributor, and an inspector. A contributed player and an inspector in a generation have the same action set in the stage game. A contributed player rewards the inspector of the previous generation, who has a different action set in his earlier life, depending on the inspector's past play. The contributed player attains high one-shot payoffs for a sufficiently long time in his later life from the contributor, which guarantees that his total average payoff can be outside V. A contributor rewards the contributed player in his earlier life, accepting possibly low payoffs. If the contributed player had deviated in his earlier life, the contributor decreases his reward. An inspector gives the appropriate payoff to the contributor, depending on whether the contributor rewarded the contributed player sufficiently. By constructing such chains of rewards and punishments lasting over generations, we prevent any deviation for each player. When the lifespans of players are sufficiently long, we can make the contributor's loss arbitrarily small, on average, and he can also attain a considerably higher average payoff in V.

[1] In finitely repeated games that do not satisfy assumption (a), Benoit and Krishna [4] proved the folk theorem by assuming that there exist multiple one-shot Nash equilibria with different payoffs. In games of overlapping generations that do not satisfy (b), Kandori [16] and Smith [24] proved several folk theorems with sufficiently long-lived players. In asynchronously repeated games that do not satisfy (c), Yoon [28] proved the folk theorem under assumptions on the payoff functions and the structure of moves. In repeated games with imperfect monitoring that do not satisfy (d), Fudenberg et al. [11] proved the folk theorem by assuming that the monitoring is public. Later, we review previous studies in detail.

In chapter 2, we study a specific two-person repeated pure coordination game of overlapping generations, with alternating moves. This model does not satisfy assumptions (b) and (c). In the pure coordination stage game, all players receive an identical payoff, given an action profile. In our stage game, two investors decide independently on whether to invest in project A or B. We assume that investors cannot obtain the profit when they invest in different projects. We further assume that the one-shot profit of investing in A by both investors is higher than that of investing in B. We consider the repeated game of overlapping generations with alternating moves, where each investor participates for finite periods. Each investor k's opponent is k－1 during the first half of his life, and then is replaced by k+1. We assume that once each investor has made his decision after entering the game, he cannot revise his decision during his life. We further assume that the one-shot profit inflates or deflates at some rate in every period. For given parameters, we characterize all subgame perfect equilibrium profiles, including the off-equilibrium paths. It is shown that the set of equilibria is discontinuous with respect to the parameters.

We first show that when the growth rates of one-shot payoffs and investors' lifespans are relatively small, all investors choose A on the path. This result holds because, in this case, the weight of investor k's payoff in his earlier life is greater than that in his later life. When investor k－1 chose A, k can receive a higher payoff by making the same choice, regardless of what k+1 chooses in the future. It is then interesting to investigate what happens if an investor chooses B by mistake. Unfortunately, it is shown that if the efficiency from choosing A rather than B is relatively small, all successive investors choose B after it has been chosen once. When k－1 chooses B, investor k can receive a higher payoff by making the same choice, regardless of what k+1 chooses. The efficient outcome can never be restored after someone chooses the inefficient action. For this result, the equilibrium with efficient outcomes in our model are not robust against players' mistakes for some range of parameters.

Second, we show that when the growth rates of one-shot payoffs and investors' lifespans are relatively large, inefficient outcomes can arise, even on the path of equilibria. In particular, the lower bound of players' (per-period) average payoffs in equilibrium can be arbitrarily small, and much smaller than the one-shot minimax payoff, when we select appropriate parameters. In this case, the weights of investors' payoffs in their later life are greater than those in their earlier life because their one-shot payoff grows exponentially over time. For each investor k, what k+1 will choose becomes more important than the choice of k－1. If investor k+1 chooses a different action to that of k, k's average payoff becomes very small because the weight of each investor's payoff in his earlier life is smaller than that in his later life. When k+1 mixes actions appropriately, conditional on k and k－1's realized actions, k becomes indifferent between A and B after any history, and obtains the payoff that is inefficient. Note that inefficient equilibria appear in the case when future payoffs increase.

In chapter 3, we consider the two-person prisoner's dilemma with finite alternating repetition. This model does not satisfy (a) or (c). In the two-person prisoner's dilemma, two players, ROW and COL, choose either cooperation (C) or defection (D), where D strongly dominates C, and the payoff from cooperation is higher than that from defection by both players. When rational players finitely repeat this game, they apply backward induction and always choose D throughout their life, in equilibria. One reasonable way to avoid such myopic actions for a player in finitely repeated games is to make him afraid that his opponent may have to imitate his previous action. That is, his opponent always cooperates if he did so in the previous period, and punishes him if not. Such a strategy profile is called a Tit-for-Tat (TFT) strategy. Kreps et al. [17] considered the symmetric two-person prisoner's dilemma with simultaneous moves, where one player has no choice but to play a TFT strategy. They showed that all sequential equilibria are almost efficient when the number of repetitions is sufficiently high. One of the other representative strategy profiles that can make payoffs in finitely repeated games efficient is the Grim-Trigger (GT) strategy. According to this strategy, the player cooperates at first, and then continues to cooperate until his opponent deviates. Once a player deviates, he continues to punish the deviator throughout his life. Fudenberg and Maskin [12] studied this strategy in general two-person games with finite repetition, and proved the folk theorem by assuming that both players have to play the GT strategy and that the number of repetitions is sufficiently high.

In our model of alternating repetition, ROW can revise his action only in odd periods, whereas COL can do so only in even periods. At the beginning of the game, nature stochastically decides whether ROW is rational and able to choose his action freely, or has to play either the GT or TFT strategy. Only ROW observes his type directly, while the probability of his type is common knowledge. We examine whether and how players can attain efficient payoffs in this model when they repeat the game for a sufficiently long time. Despite the negative results in several studies (see chapter 3), which show that inefficient outcomes that never arise in simultaneous move games sometimes cannot be excluded from the set of equilibrium payoffs in asynchronous move games, we show that *all* sequential equilibria remain almost efficient, even though the structure of repetition in our model is asynchronous.

The intuitive reason for this result is as follows. First, it is optimal for COL to always choose C, except in the last stage if ROW definitely plays either the Grim-Trigger or Tit-for-Tat strategy. Second, if rational ROW chooses D when COL chooses C, ROW's rationality becomes common knowledge, and players continue to choose D in all subsequent stages, yielding low continuation payoffs. However, if rational ROW chooses C when COL does the same, COL cannot determine ROW's type, and the players can obtain higher continuation payoffs from cooperation. Therefore, it is optimal for COL to cooperate, except in the last few stages if rational ROW also cooperates. Because ROW rationally expects this, he also cooperates, except in the last few stages if COL does so as well, which guarantees the existence of cooperative equilibria with efficient payoffs.

In chapter 4, we consider the two-person repeated prisoner's dilemma with private monitoring. This model does not satisfy (d). At the end of every period, each player receives a noisy private signal with a certain probability, depending on his opponent's current action instead of observing it directly. We assume that the probability distribution of private signals changes over time, depending on calendar dates. We also admit the case where the distribution of signals in any period is different from that in any other period. It is proved that efficient sequential equilibria exist with sufficiently patient players when the monitoring is almost perfect in every period. In order to prove this, we construct the strategy called the Markov strategy, which only depends on each player's own action, the private signal in the previous period, and the calendar date. We observe that when a player's opponent adopts a Markov strategy, he is indifferent between C and D in every period and, therefore, cannot change his continuation payoff by himself. When players are sufficiently patient and the monitoring technology is almost perfect in every period, such a strategy exists for the payoff vector between the payoff of (C, C) and (D, D), which ensures that efficient payoffs are attainable in this game. This argument also applies to the game with alternating repetition, which does not satisfy (c).

Our model in chapter 1 can be used to examine the problem of the pension system in Japan. Today, in Japan, retired elderly people can receive relatively high pensions, compared with insurance premiums they paid in their younger days. On the other hand, the working generation will have to accept low pensions, even though they pay high insurance premiums today. Retired people and the working generation can be regarded in our model as contributed players with high payoffs and contributors with low payoffs, respectively. Our result shows that such a phenomenon can arise, not because of the failure of the system, but *on the path of equilibria*.

Our model in chapter 2 allows one-shot payoffs to keep increasing in every period, which seems natural because the real-world economy does grow over time. Interestingly, our result shows that inefficient outcomes may appear in this case, rather in one of deflation.

The two-person prisoner's dilemma considered in chapters 3 and 4 can be viewed as binary price competition by duopolists[2]. In incomplete competition, including in a duopoly, we sometimes observe price rigidity. In such a case, we can analyze the model more appropriately by assuming that the structure of repetition is asynchronous. Our model can be used to investigate how the ease of forming the coalition changes when price rigidity is given exogenously. Our results in chapter 3 and in section 3 of chapter 4 show that the ease of forming the coalition does not much depend on price flexibility.

Our model in chapter 4, with a varying structure of monitoring over time, includes the cases of technological improvements in monitoring, small increases in uncertainty over time, periodic changes in the distribution of signals. Our result shows that duopolists can collude, even in such circumstances.

[2] Two duopolists choose either a high ($p(h)$) or a low ($p(\ell)$) price. The total demand when the lower price is $p(k)$ is $d(k)$. The marginal cost is c. When they choose the same price, they share the demand evenly. When they choose different prices, the player who chooses the lower price monopolizes the demand, whereas his opponent gets nothing. Assume $p(h)>p(\ell)>c>0$ and $d(\ell)>d(h)>0$. When $0<d(\ell)(p(\ell)-c)<d(h)(p(h)-c)<2d(\ell)(p(\ell)-c)<2d(h)(p(h)-c)$ holds, then this model is a standard two-person prisoner's dilemma. Indeed, $d(\ell)=2$, $d(h)=1$, $p(\ell)=2$, $p(h)=4$, and $c=1$ satisfy these inequalities.

Chapter 1

Extension of the Equilibrium Payoff Set in Overlapping Generations Games

1.1 Introduction

In repeated games, players can obtain payoffs that are not obtainable in one-shot games. Under the condition of full dimensionality, Fudenberg and Maskin [12] proved that any payoffs in V, the set of feasible and individually rational payoffs of the stage game, are attainable as average subgame perfect equilibrium payoffs in repeated games. Many studies, including theirs, assume that the number of repetitions is infinite. However, such an assumption may not reflect real-world conditions. For example, companies continue to exist in the distant future, while they are controlled by employees whose tenures are finite. To analyze such an economic model, overlapping generations (OLG) games are introduced. In OLG games, players in the same generation interact for a sufficiently long time, and then are gradually replaced by successors in the next generation. Kandori [16] was the first to prove the folk theorem in general N-person OLG games.

In the versions of folk theorems in OLG games proved by Kandori [16], and in other subsequent studies, it is required that any (per-period) average equilibrium payoffs are included in the feasible and individually rational one-shot payoff set V. In contrast, in our model of two-person OLG games, we show that there exist subgame perfect equilibria where some players can attain possibly high average payoffs outside V. We also guarantee that players who give high one-shot payoffs to other players do not have to decrease their payoffs, on average, when their lifespan is sufficiently long. Consequently, our result consequently establishes the new folk theorem, which differs from the results of previous studies.

In the classical repeated games with fixed players, the players cannot attain payoffs outside V in equilibria because even if a player desires to do so, his opponent can refuse it by taking the one-shot best response against him, and his one-shot gain decreases. Therefore, we must consider how to punish the player who failed to contribute to other players, who should have obtained high payoffs outside V. In OLG games, a player who obtained low payoffs outside V in his earlier life in order to contribute to another player can be rewarded before his retirement by the other younger player. In our model of two-person OLG games, players in each generation are classified into three types, each with different roles: a contributed player, a contributor, and an inspector. A contributed player and an inspector of each generation have the same action set for the stage game. A contributed player rewards the inspector of the previous generation in his earlier life, depending on his past choices. He attains high one-shot payoffs for a sufficiently long time in his later life from the contributor, which guarantees that his total average payoff can be outside V. A contributor rewards the contributed player in his earlier life, accepting possibly low payoffs. If the contributed player deviated in his earlier life, the contributor decreases his reward. An inspector gives the appropriate payoff to the contributor, depending on whether or not the contributor rewarded the contributed player sufficiently. By constructing such chains of rewards and punishments lasting over generations, we can prevent any deviation for each player. When the lifespans of players are sufficiently long, we can make the contributor's loss arbitrarily small, on average, and he can also attain the considerably high average payoff in V.

Several variations of folk theorems in OLG games have been proved. The folk theorem in two-person OLG games was first proved by Salant [23], under certain restrictions on payoff functions of the stage game. Following Kandori [16], Smith [24] proved several versions of folk theorems, with stronger results, in general N-person OLG games. Gossner [13] showed that the folk theorem holds in OLG games without public randomizations. These studies analyzed equilibria where different generations who have identical action sets play the same strategy and get the same payoff. Therefore, any pair of players' average payoffs must be in V. In our model, on the other hand, the contributed player can attain high payoffs outside V at the expense of the contributor's low payoff. In order to reward the contributor before his retirement, we must assume the existence of an inspector, who has the same action set as that of the contributed player, but plays a completely different strategy.

Our model can be seen as a class of repeated games with long-run and short-run players, because players' lifespans differ among the types. A representative study that analyzed this class of repeated games is that of Fudenberg et al. [9], who characterized the set of subgame perfect equilibria. In their model, they assume that short-run players take part in the game only for one period, whereas long-run players exist for an infinitely long time. They analyzed the set of equilibrium payoffs for long-run players only. Our model, in contrast, assumes that all players' lifespans are finite and characterizes the set of average equilibrium payoffs for all players.

In the remainder of this chapter, we define the model and prove the above result.

1.2 The Model and Result

Let G be a two-person stage game. Let $S = \{1,\ 2\}$ be a set of players in G. Let A_i be the pure action set of $i \in S$ and denote $A = \prod_{i\in S} A_i$. Denote i's one-shot payoff function as $g_i\colon\ A \to R$. The minimax profile of $j \neq i$ against i and his best response are respectively defined as follows:

$$m_j^i \in \arg\min_{a_j\in A_j} \max_{a_i\in A_i} g_i(a),$$
$$m_i^i \in \arg\max_{a_i\in A_i} g_i(m_j^i,\ a_i).$$

We assume w.l.o.g. that $\forall i\ g_i(m^i) = 0$ holds. The set of feasible and strongly rational payoffs are defined as $V = co(g(A)) \cap R^{2++}$. We assume that $\dim V \geq 1$. Define i's best payoff, worst payoff, range of payoffs, best one-shot gain by the deviation, and players' payoffs when minimaxing i as follows, respectively:

$$v_i^* = \max_{a\in A} g_i(a),$$
$$\underline{v_i} = \min_{a\in A} g_i(a),$$
$$\Delta v_i = v_i^* - \underline{v_i},$$
$$d_i = \max_{a\in A,\ a_i'\in A_i}(g_i(a_j,\ a_i') - g_i(a)),$$
$$z^i = g(m^i).$$

For $i \in S$, denote $K(1)$, $K(2)$, $K(3)$, and T as natural numbers.
The overlapping generations game $OLG(G;\ K(1),\ K(2),\ K(3),\ T)$ with perfect monitoring is defined as follows (also see figure 1 on the next page).

·For $i \in S$ of each generation, there exist three types of players: a contributed player ($Hi(i)$), a contributor ($Cont(j)$), and an inspector ($Ins(i)$). $Hi(i)$ and $Ins(i)$ have the action set A_i, whereas $Cont(j)$ has A_j.

·Each generation of $Hi(i)$, except the first generation, lives for $K(2)+K(3)$ periods, whereas the first generation, which is born at period 1, lives for $K(3)$ periods. His opponent during his earlier $K(2)$ periods is $Ins(j)$ of the previous generation, whose action set is A_j. His opponent during his later $K(3)$ periods is $Cont(j)$.

·$Cont(j)$ lives for $K(3)+T+K(1)$ periods. His opponent during his earlier $K(3)$ periods is $Hi(i)$. His opponent during his later $T + K(1)$ periods is $Ins(i)$.

·$Ins(i)$ lives for $T + K(1) + K(2)$ periods. His opponent during his earlier $T + K(1)$ periods is $Cont(j)$. His opponent during his later $K(2)$ periods is $Hi(j)$ of the next generation, whose action set is A_j.

We use a public randomizing device in each period. Each player's total payoff is discounted by the length of his life, which we call his average payoff. The following result holds.

(Folk theorem with average payoffs outside V)
Suppose $v(Hi(i)) \in (0,\ v_i^*)$ and $v(ji) = (v(Cont(j)),\ v(Ins(i))) \in V$ hold for $i \in S$.
Then, $\forall\epsilon > 0\ \exists\underline{K}(1)\ \forall K(1) \geq \underline{K}(1)\ \exists\underline{K}(2)\ \forall K(2) \geq \underline{K}(2)\ \exists\underline{K}(3)\ \forall K(3) \geq \underline{K}(3)\ \exists\underline{T}\ \forall T \geq \underline{T}$,
and there exists a subgame perfect equilibrium in $OLG(G;\ K(1),\ K(2),\ K(3),\ T)$,
where the distances of $Hi(i)$, $Cont(j)$, and $Ins(i)$'s average payoffs in equilibrium
from $v(Hi(i))$ and $v(ji)$ are less than ϵ.
We admit that $(v(Hi(i)),\ v(Cont(j))) \notin V$.

Proof. Choose each $v(Hi(i)) \in (0,\ v_i^*)$ and $v(ji) = (v(Cont(j)),\ v(Ins(i))) \in V$ for $i \in S$, and $\epsilon > 0$. Define the variables as follows:

$$u^{Cont(i)} \in V \text{ satisfies } u_i^{Cont(i)} < v(Cont(i)),$$
$$\underline{u_i} = \min\{v(Ins(i)),\ u_i^{Cont(i)},\ u_i^{Cont(j)}\},$$
$$\underline{u} = \min_{i\in S} u_i,$$
$$\underline{d} = \min_{i\in S} d_i,$$

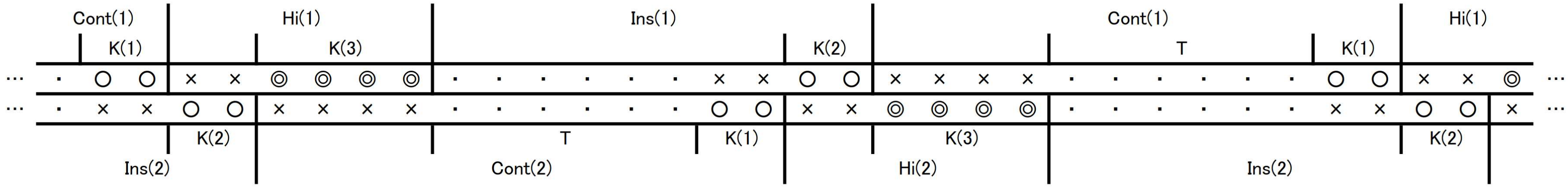

Figure 1: The structure of overlapping generations

$$M \text{ satisfies } M\underline{u} > \underline{d},$$
$$\overline{d} = \max_{i\in S} d_i,$$
$$\Delta > \overline{d} + (M-1)\max_{i\in S,\ j\neq i}(|z_i^j| + \underline{u_i}),$$
$$\underline{v^*} = \min_{i\in S} v_i^*,$$
$$\overline{d} = \max_{i\in S} d_i,$$
$$K(1) \text{ satisfies } K(1)\underline{v^*} > \overline{d} + \Delta,$$
$$\underline{v} = \min_{i\in S} \underline{v_i},$$
$$K(2;\ 1) \text{ satisfies } K(1)\underline{v} + K(2;\ 1)\underline{v^*} > K(1)\underline{v^*},$$
$$K(2;\ 2) \text{ satisfies } K(2;\ 2)\underline{v^*} > \Delta,$$
$$K(2) = K(2;\ 1) + K(2;\ 2),$$
$$K(3) \text{ satisfies } K(2)\underline{v} + K(3)v(Hi(i)) > (K(2)+K(3))(v(Hi(i)) - \epsilon),$$
$$K(2)\overline{v} + K(3)v(Hi(i)) < (K(2)+K(3))(v(Hi(i)) + \epsilon) \text{ for } i \in S, \text{ and}$$
$$K(2)\underline{v} + K(3)\min_{i\in S} v(Hi(i)) > K(2)\overline{v},$$
$$T \text{ satisfies } K(3)\underline{v} + Tv(Cont(i)) + K(1)\underline{v} > (K(3)+T+K(1))(v(Cont(i)) - \epsilon),$$
$$K(3)\overline{v} + Tv(Cont(i)) + K(1)\overline{v} < (K(3)+T+K(1))(v(Cont(i)) + \epsilon) \text{ for } i \in S,$$
$$K(3)\underline{v} + T\min_{i\in S} v(Cont(i)) > K(3)\overline{v} + T\max_{i\in S} u_i^{Cont(i)},$$
$$Tv(Ins(i)) + (K(1)+K(2))\underline{v} > (T+K(1)+K(2))(v(Ins(i)) - \epsilon), \text{ and}$$
$$Tv(Ins(i)) + (K(1)+K(2))\overline{v} < (T+K(1)+K(2))(v(Ins(i)) + \epsilon) \text{ for } i \in S,$$
$$w_i^{Normal} = w_i^{Early(i)} = d_i,$$
$$w_i^{Late(i)} = 0,$$
$$w_i^{Early(j)} = w_i^{Late(j)} = d_i + (M-1)(|z_i^j| + \underline{u_i}),$$
$$H(1) \in \{Normal,\ Early(1),\ Early(2),\ Late(1),\ Late(2)\},$$
$$H(2) \in \{Good,\ Bad\},$$
$$H(3) \in \{Good,\ Bad\},$$
$$a(1,\ i;\ H(1)) \text{: the profile with one-shot payoff } \frac{w_i^{H(1)}}{K(1)} \text{ to } Cont(i),$$
$$f_j \in (K(1)v_j^*,\ K(1)\underline{v_j} + K(2;\ 1)v_j^*),$$
$$x_j \in (K(1)\underline{v_j},\ K(1)\overline{v_j}),$$
$$a(2,\ i;\ x_j) \text{: the profile with one-shot payoff } \frac{1}{K(2;\ 1)}(f_j - x_j) \text{ to } Ins(j),$$
$$a(2,\ i;\ H(1)) \text{: the profile with one-shot payoff } \frac{w_j^{H(1)}}{K(2;\ 2)} \text{ to } Ins(j),$$
$$a(3,\ i;\ Good) \text{: the profile with one-shot payoff } v(Hi(i)) \text{ to } Hi(i),$$
$$a(3,\ i;\ Bad) \text{: the profile with one-shot payoff 0 to } Hi(i),$$
$$a(4,\ i;\ Good) \text{: the profile with one-shot payoffs } v(ji) \text{ to } Cont(j) \text{ and } Ins(i),$$
$$a(4,\ i;\ Bad) \text{: the profile with one-shot payoffs } u^{Cont(j)} \text{ to } Cont(j) \text{ and } Ins(i),$$
$$e \text{: the one-shot Nash profile.}$$

Players' strategy profiles are defined by the following phases:
The game starts with $(K(3),\ Good)$ at period 1.

·Phase $(K(1);\ H(1))$:
This phase begins $K(1)$ periods before $Cont(i)$ dies and lasts for $K(1)$ periods.
Players $Cont(i)$ and $Ins(j)$ play $a(1,\ i;\ H(1))$.
If $Ins(j)$ deviates, then continue to play e.
At the end of this phase, set x equal to the total payoff $Ins(j)$ gained during this phase.

·Phase $(K(2;\ 1);\ x)$:
This phase begins just after $Cont(i)$ died and lasts for $K(2;\ 1)$ periods.
Players $Ins(j)$ and $Hi(i)$ play $a(2,\ i;\ x)$.
If $Hi(i)$ deviates, then continue to play e until $Ins(j)$ dies and set $H(2) =$ "*Bad*."

·Phase $(K(2;\ 2);\ H(1))$:
This phase begins $K(2;\ 1)$ periods after $Cont(i)$ died and lasts for $K(2;\ 2)$ periods.
Players $Ins(j)$ and $Hi(i)$ play $a(2,\ i;\ H(1))$.
If $Hi(i)$ deviates, then continue to play e and set $H(2) =$ "*Bad*."
At the end of this phase, set $H(2) =$ "*Good*" if there is no deviation.

·Phase $(K(3);\ H(2))$:
This phase begins just after $Ins(j)$ died and lasts for $K(3)$ periods.
Players $Hi(i)$ and $Cont(j)$ play $a(3,\ i;\ H(2))$.
If $Cont(j)$ deviates, then continue to play e until $Hi(i)$ dies and set $H(3) =$ "*Bad*" and $H(1) =$ "*Good*."
At the end of this phase, set $H(3) =$ "*Good*" and $H(1) =$ "*Good*" if there is no deviation.
Then go to phase $(T;\ H(3))$.

Set $t = 1$. t increments as the stage game is played.
·Phase $(T;\ H(3))$:
Players $Cont(j)$ and $Ins(i)$ play $a(4,\ i;\ H(3))$ as long as $t \leq T$.
Set $H(1) =$ "$Early(k)$" and go to $(T;\ Early(k))$ if k deviates at $t \leq T$.
Set $H(1) =$ "$Late(k)$" and go to $(T;\ Late(k))$ if k deviates at $t \geq T - M + 1$.

·Phase $(T;\ Early(k))$:
Play m^k for M periods.
Set $H(1) =$ "$Early(\ell)$" and go to $(T;\ Early(\ell))$ if $\ell \neq k$ deviates at $t \leq T - M$.
Set $H(1) =$ "$Late(\ell)$" and go to $(T;\ Late(\ell))$ if $\ell \neq k$ deviates at $t \geq T - M + 1$.
Go back to $(T;\ H(3))$ when there is no deviation.

·Phase $(T;\ Late(k))$:
Play m^k as long as $t \leq T$.
Set $H(1) =$ "$Late(\ell)$" and go to $(T;\ Late(\ell))$ if $\ell \neq k$ deviates.

$Hi(i)$'s total payoff lies on $[K(2)\underline{v_i} + K(3)v(Hi(i)),\ K(2)\overline{v_i} + K(3)v(Hi(i))]$.
$Cont(j)$'s total payoff lies on $[(K(3) + K(1))\underline{v_j} + Tv(Cont(j)),\ (K(3) + K(1))\overline{v_j} + Tv(Cont(j))]$.
$Ins(i)$'s total payoff lies on $[Tv(Ins(i)) + (K(1) + K(2))\underline{v_i},\ Tv(Ins(i)) + (K(1) + K(2))\overline{v_i}]$.
The difference between their average payoffs in equilibrium and the target payoffs are at most ϵ.

In the remainder of this section, we prove that each player's strategy is optimal. For each phase, we consider the "worst-case scenario," where the incentive to deviate is greatest.

·$Ins(j)$'s deviation from $(K(1);\ H(1))$:
$Ins(j)$'s total continuation payoff during $K(1) + K(2;\ 1)$ periods remains unchanged at f_j, regardless of his choice during this phase.

·$Hi(i)$'s deviation from $(K(2;\ 1);\ x)$ or $(K(2;\ 2);\ H(1))$:
$Hi(i)$'s one-shot continuation payoff during $(K(3);\ H(2))$ falls from $v(Hi(i))$ to 0 if he deviates during these phases.

·$Cont(j)$'s deviation from $(K(3);\ H(2))$:
$Cont(j)$'s one-shot continuation payoff during $(T;\ H(3))$ falls from $v(Cont(j))$ to $u_j^{Cont(j)}$ if he deviates during this phase.

·Player $\ell \neq k$'s deviation from $(T;\ \cdot)$ at $t \leq T - M$:
Suppose that $Cont(\ell)$ deviates from $(T;\ Early(k))$ with $H(3) =$ "Bad."
If he does not deviate, he gets at least $Mz_\ell^k + u_\ell^{Cont(\ell)} + w_\ell^{Early(k)}$.
If he deviates, he gets at most $z_\ell^k + d_\ell + w_\ell^{Early(\ell)}$.
The former is greater than the latter by the definition of M, $w_\ell^{Early(k)}$ and $w_\ell^{Early(\ell)}$.

·Player $\ell \neq k$'s deviation from $(T;\ \cdot)$ at $t \geq T - M + 1$:
Suppose that $Cont(\ell)$ deviates from $(T;\ Late(k))$.
If he does not deviate, he gets at least $Mz_\ell^k + w_\ell^{Late(k)}$.
If he deviates, he gets at most $z_\ell^k + d_\ell + w_\ell^{Late(\ell)}$.
The former is greater than the latter by the definition of M, $w_\ell^{Late(k)}$ and $w_\ell^{Late(\ell)}$.

Therefore, players' strategy profiles prevent any unilateral deviation and form a subgame perfect equilibrium. $\square$

Chapter 2

The Discontinuous Inefficiency and Unstability of Equilibria in Alternately Repeated Pure Coordination Games with Overlapping Generations of Investors

2.1 Introduction

In repeated games, players can obtain payoffs that are not obtainable in one-shot games. Under the condition of full dimensionality, Fudenberg and Maskin [12] showed that any feasible and individually rational payoffs of the stage game are attainable as subgame perfect equilibrium payoffs in repeated games. Many studies, including theirs, assume that the repetition is simultaneous. This means there is no period where some players can revise their actions while the others cannot do so. However, in the real world, the repetition is sometimes not simultaneous. For example, stock markets in Japan and the United States do not operate at the same time, just because they are located in different time zones.

Motivated by this issue, we study a specific two-person repeated pure coordination game of overlapping generations, with alternating moves. In the stage game of pure coordination, all players receive the identical payoff, given an action profile. In our stage game, two investors decide independently on a project, A or B, in which to invest. We assume that investors cannot obtain the profit when they invest in different projects. We also assume that the one-shot profit of investing in A by both investors is higher than that of investing in B. We consider the repeated game of overlapping generations with alternating moves, where each investor participates for finite periods of time. Each investor k's opponent is $k-1$ during the first half of his life, and then is replaced by $k+1$. We assume that once each investor has made his decision upon entering the game, he cannot later revise his decision. We also assume that the one-shot profit inflates or deflates at some rate in every period. For given parameters, we characterize all subgame perfect equilibrium profiles, including the off-equilibrium paths. It is shown that the set of equilibria is discontinuous with respect to the parameters.

We first show that when the growth rates of one-shot payoffs and investors' lifespans are relatively small, all investors choose A on the path. This result holds because each investor k's weight of payoff in his earlier life is greater than in his later life. When investor $k-1$ chooses A, k can get a higher payoff by making the same choice, regardless of what $k+1$ chooses in the future. It is then interesting to investigate what happens if an investor chose B by mistake. Unfortunately, it is shown that if the efficiency from choosing A rather than B is relatively small, all successive investors choose B after it has been chosen once. When $k-1$ chooses B, investor k can get a higher payoff by making the same choice, regardless of what $k+1$ chooses. The efficient outcome can never be restored after someone chooses the inefficient action. For this result, the equilibrium with efficient outcomes in our model is not robust against a player's mistake, for some range of parameters.

Second, we show that when the growth rates of one-shot payoffs and investors' lifespans are relatively large, inefficient outcomes can arise, even on the path of equilibria. In particular, the lower bound of players' (per-period) average payoffs in equilibrium can be arbitrarily small, and much smaller than the one-shot minimax payoff, when we use appropriate parameters. In this case, investors' payoff weights in their later life are greater than in their earlier life, because their one-shot payoff grows exponentially over time. For each investor k, what $k+1$ will choose becomes more important than what $k-1$ chooses. If investor $k+1$ chooses a different action to that of k, k's average payoff becomes very small because each investor's weight of payoffs in his earlier life is smaller than that in his later life. When $k+1$ mixes actions appropriately, conditional on k and $k-1$'s realized actions, k becomes indifferent between A and B after any history, and obtains an inefficient payoff. Note that inefficient equilibria appear in the case when payoffs in the future grow.

How the set of subgame perfect equilibria differs between synchronously and asynchronously repeated games has been researched for a long time. The pure coordination game is a class of stage games in which the set of subgame perfect equilibria can vary, depending on the structure of the repetition. When a pure coordination game has multiple one-shot Nash equilibria with different payoffs, inefficient payoffs cannot be excluded from the set of subgame perfect equilibria in the corresponding repeated game with simultaneous repetition. When the repetition is alternating, however, each player can get an efficient one-shot payoff by choosing the efficient action when his opponent does so as well, which implies that inefficient outcomes can be avoided.

In alternately repeated pure coordination games, Lagunoff and Matsui [18] proved that only the most efficient payoff is attainable as a subgame perfect equilibrium payoff when players are sufficiently patient. In their model, one player can revise his action only in odd periods, whereas his opponent can only do so in even periods. Both players' continuation payoffs are identical, for any strategy and history. Thus, the subgame perfection guarantees that each player's infimum of continuation payoff when his opponent chooses the efficient action is at least the discounted sum of the one-shot payoff by choosing the efficient action and his opponent's infimum of continuation payoff when he chooses the efficient action. Therefore, the continuation payoffs when one player chooses the efficient action are always Pareto-optimal. Dutta [6] applied this result to finitely asynchronous repeated games of common interest, where the stage game has a unique Pareto-optimal payoff vector. He showed that only Pareto-optimal payoffs can be attained as average payoffs of subgame perfect equilibria. Because the number of repetitions is finite, each player can consider final periods, where the players who cannot move are choosing Pareto-optimal actions. In such periods, choosing the most efficient action is optimal for each player who can move. Applying backward induction, we show that only Pareto-optimal actions are played throughout the equilibrium path.

In asynchronously repeated games of common interest, players do not want to punish their opponents because their own payoffs may also decrease as a result. Yoon [28] pointed out this peculiarity, and proved the folk theorem for infinitely repeated games, including asynchronously repeated games under conditions of non-equivalent utility and finite periods of inaction. In his thesis, the infimum of individually rational payoffs supported by the folk theorem is the minimax value of the stage game. Takahashi and Wen [25] considered a specific three-person asynchronously repeated game, and showed that the minimum of average subgame perfect equilibrium payoffs for a player is less than his minimax value in the stage game. In their model, some players can move in every period, while the other cannot do so. When the former chooses appropriate actions, they make the latter player's one-shot payoff equal to the maxmini value of the stage game during periods when he cannot move. This example clarified that the lower bound of subgame perfect equilibrium payoffs in asynchronously repeated games crucially depends on the structure of the repetition. Wen [26] and Yoon [29] defined the effective minimax value of asynchronously repeated games, which takes the structure of the repetition into account, and proved the folk theorem in which the infimum of individually rational payoffs is equal to this value.

In our model, all players are assumed to live a finite life. When players reach the end of the game, efficient outcomes are not always included in the set of subgame perfect equilibrium payoffs, because each player always chooses a myopic behavior at the end of his life. In finitely repeated games where players quit the game at the same time, whether players can attain efficient payoffs in a subgame perfect equilibrium depends on their utility functions in the stage game. Benoit and Krishna [4] proved the folk theorem in finitely repeated games by assuming that there exist multiple one-shot Nash equilibria with different payoffs. Another class of repeated games where players do not live forever is the overlapping generations game, which we study in this chapter. Kandori [16] and Smith [24] proved folk theorems in overlapping generations games by assuming that players' lifespans are sufficiently long and that they are sufficiently patient.

Our results are related to those of previous studies as follows. In contrast to Lagunoff and Matsui [18], players' lifespans are finite in our model, and each player enters the game in a different period. Therefore, players' discounted payoffs in each period are not identical, because their one-shot payoffs are multiplied by different discount rates. We cannot regard one player's continuation payoff as the same as that of his opponent's, and each player's continuation payoff must be calculated independently. Moreover, we can also consider the case when one-shot payoffs grow at some rate, which Lagunoff and Matsui [18] were not able to do because players' lives are infinite in their model. Lagunoff and Matsui [18] pointed out that if the overtaking criterion is employed to calculate payoffs, their efficient result may not hold because payoffs in the future do not decrease. Our inefficient result also appears because each player's payoff in the future can have more weight than his payoff today. Hence, in order for repeated alternating move games to exclude inefficient outcomes, it is important that players' future payoffs strictly decrease.

Our result is different from that of Dutta [6] in that inefficient outcomes can be realized in our model, even though each player's lifespan is finite. In alternately repeated pure coordination games, where all players die at the same time, players always choose efficient actions in equilibria, applying backward induction. In our overlapping generations game, on the other hand, when player k chooses the efficient action, his successor $k+1$ can decrease his continuation payoff by choosing the inefficient action, with a higher probability. Because the continuation payoff does not have to decrease throughout each k's life, this punishment by $k+1$ is possible. Considering chains of such punishments over generations, we can construct inefficient equilibria in our overlapping generations game.

The model considered in Takahashi and Wen [25] is a three-person repeated game with infinitely lived players, and is not a game of pure coordination. In their game, punishing only one player without decreasing the punishers' payoffs is possible, using the structure of asynchronous repetition. They also assumed that mixed actions are observable, and that a player who cannot move in a period will have to play the same mixed action chosen in the previous period. Our model, in contrast, is a two-person repeated pure coordination game with unobservable mixed actions, where the maxmini value and the minimax value of the stage game are identical.

In what follows, we define the model and prove each result stated above.

2.2 The Model and Results

2.2.1 Definition of The Model

The stage game of pure coordination in our model is defined as follows. Two players choose one of two actions, A or B. Their payoff is 1 when both choose A, and $\alpha \in (0, 1)$ when both choose B. Their payoff is 0 when different actions are chosen. This stage game has three Nash equilibria. The pure equilibria are (A, A) with payoff 1 and (B, B) with payoff α. There exists one mixed equilibrium where they choose A with probability $\alpha/(1+\alpha)$, with payoff $\alpha/(1+\alpha)$, which is the same as the minimax and maxmin value. When the repetition is simultaneous, the lower bound of the subgame perfect equilibrium payoff is this minimax value. When the repetition is alternating and players live forever or die at the same time, only A is played in the path of all subgame perfect equilibria, given that they are patient, as Lagunoff and Matsui [18] and Dutta [6] showed.

The overlapping generations game with alternating repetition $G(\alpha, T, R)$ is defined as follows. Each player $k \geq 1$ takes part in the game for $2T$ periods, during the period $(k-1)T+1$ and $(k+1)T$. The opponent of player $k \geq 1$ during his first T periods is $k-1$, and is then replaced by $k+1$. Exclusively, the initial player 0 takes part in the game from period 0 to T and his payoff at period 0 is assumed to be 0. We assume that each player cannot revise the action he took at the beginning of his participation. We also assume that one-shot payoffs grow or shrink in each period at some rate $R>0$. At period $s \geq 1$, the participating players' one-shot payoffs are multiplied by R^{s-1}. For example, their payoff is αR^{s-1} when (B, B) is played at period s. Players can perfectly observe their past. There is no public randomizing device available and the realized pure action is observable. Note that when $R>1$ (resp. $R<1$), then R^T tends to ∞ (resp. 0) as T becomes large.

Let $U=1+R+\cdots+R^{T-1}$. The payoff of player $k \geq 1$ is equal to $R^{(k-1)T}(1+R^T)U$ when only (A, A) is played throughout his life. When the total payoff of $k \geq 1$ is divided by this value, it is called his average payoff. When his total payoff is divided by $R^{(k-1)T}$, it is called his discounted payoff. For example, a player's average payoff and discounted payoff are 1 and $(1+R^T)U$, respectively, when (A, A) is chosen throughout his life. For $a \in \{A, B\}$, let $H(a)$ be the set of histories where a player enters the game and his predecessor is taking a. Let H be the set of all histories when some player enters the game. For $h \in H$, let $h \circ a \in H(a)$ be the history where the player entering immediately after h chooses a.

2.2.2 The Results

From now on, we characterize players' behavior, both on and off the path, in subgame perfect equilibria of $G(\alpha, T, R)$ as the parameters vary (also see figure 2 on the next page).

1. When $R^T < 1/\alpha$, then for all subgame perfect equilibria, players choose A after $h \in H(A)$.

Proof. Suppose that a player's predecessor chooses A. When he does so, his discounted payoff is at least U. When he chooses B, on the other hand, it is at most $R^T \alpha U$. The former is strictly greater than the latter. □

2. When $\alpha/(1-\alpha) < R^T < 1/\alpha$, then for all subgame perfect equilibria, players choose A after any history $h \in H$.

Proof. Suppose that a player's predecessor chooses B. When the player chooses A, his discounted payoff is at least $R^T U$, from 1. When he chooses B, on the other hand, it is at most $(1+R^T)\alpha U$. The former is strictly greater than the latter. □

3. When $R^T < \alpha$, then for all subgame perfect equilibria, players choose B after $h \in H(B)$.

Proof. Suppose that a player's predecessor chooses B. When the player chooses B, his discounted payoff is at least αU. When chooses A, on the other hand, it is always $R^T U$, from 1. The former is strictly greater than the latter. □

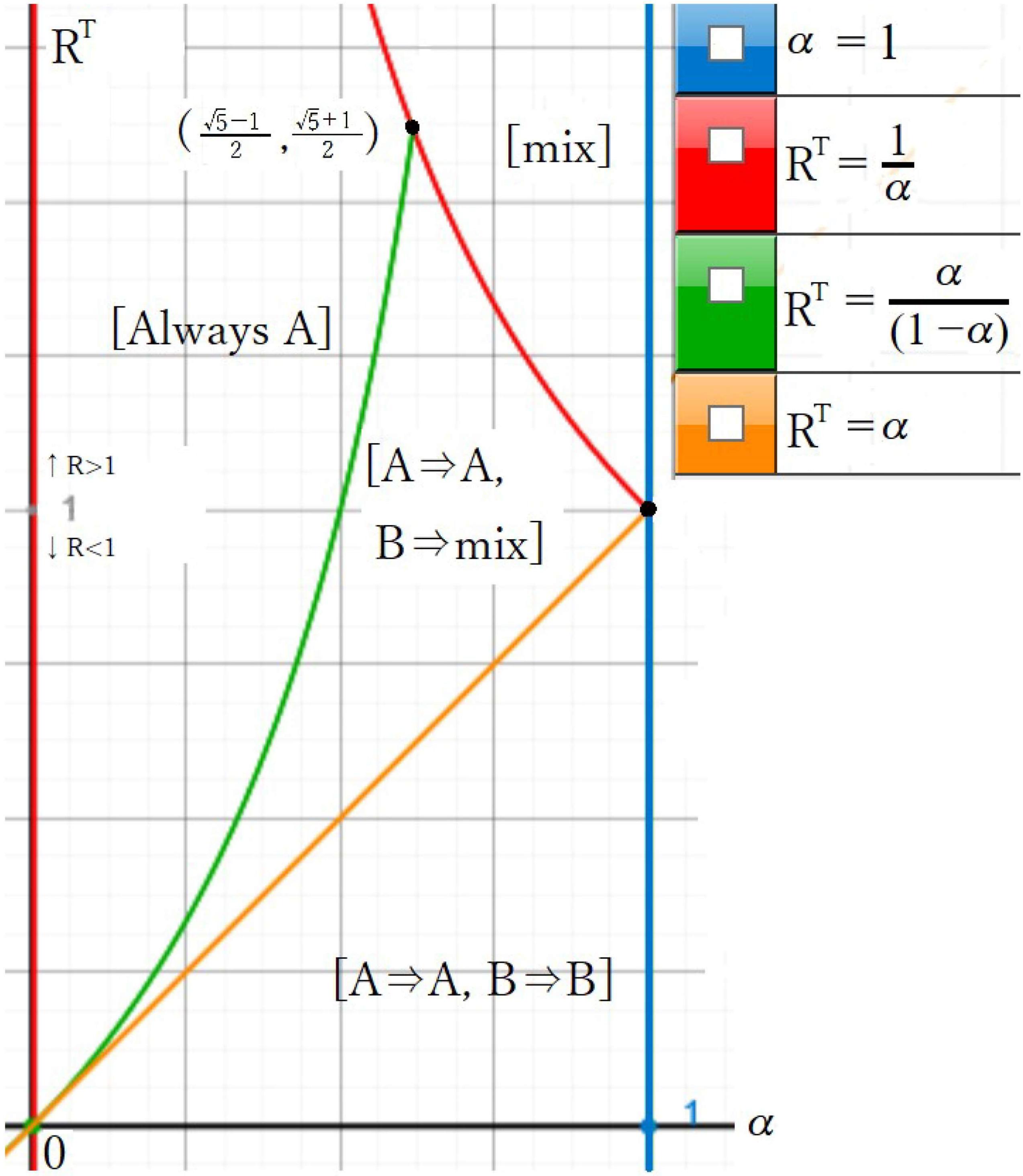

Figure 2: The equilibrium strategy profiles

4. When $\alpha < R^T < \min\{1/\alpha,\ \alpha/(1-\alpha)\}$, then for all subgame perfect equilibria, players k and $k+1$ choose one of the following actions after $h \in H(B)$.

4.I. When k chooses B with probability $r \in (0,\ 1)$ after h, then $k+1$ chooses B with probability $(R^T - \alpha)/\alpha R^T$ after $h \circ B$.

4.II. When k chooses B with probability one after h, then $k+1$ chooses B with probability $p \in [(R^T - \alpha)/\alpha R^T,\ 1]$ after $h \circ B$.

4.III. When k chooses A with probability one after h, then $k+1$ chooses B with probability $p \in [0,\ (R^T - \alpha)/\alpha R^T]$ after $h \circ B$.

Proof of 4.I. Because k is indifferent between A and B, $R^T U = \alpha U + p R^T \alpha U$ must hold when $k+1$ chooses B with probability p after $h \circ B$. Then, $p = (R^T - \alpha)/\alpha R^T$. The other results are shown analogously. □

By 1 and 3, the following result holds.
5. (The unstability of equilibria): $\forall \alpha \in (0,\ 1)$, $\forall R \in (0,\ 1)$, $\exists T(0)$, $\forall T \geq T(0)$, $\forall h \in H(B)$, and in all subgame perfect equilibria, players choose B after h.

Proof. Take α and $R<1$. When T is sufficiently large, $R^T < \alpha$ holds, which guarantees that players choose B with probability one after $h \in H(B)$ by 3. □

Next, we show that inefficient equilibria can arise under some range of parameters.

6. When $R^T > 1/\alpha$, then for all subgame perfect equilibria, players k and $k+1$ choose one of the following actions after $h \in H(A)$.

6.I. When k chooses A with probability $r \in (0,\ 1)$ after h, then $k+1$ chooses B with probability $q \in [1/\alpha R^T,\ 1]$ after $h \circ B$ and chooses A with probability $\alpha \mathrm{q} - (1/R^T)$ after $h \circ A$.

6.II. When k chooses A with probability one after h, then $k+1$ chooses B with probability $q \in [0,\ 1]$ after $h \circ B$ and chooses A with probability $p \in [\max\{0,\ \alpha q - (1/R^T)\},\ 1]$ after $h \circ A$.

6.III. When k chooses B with probability one after h, then $k+1$ chooses B with probability $q \in [1/\alpha R^T,\ 1]$ after $h \circ B$ and chooses A with probability $p \in [0,\ \alpha q - (1/R^T)]$ after $h \circ A$.

Proof of 6.I. Because k is indifferent between A and B, $U + \mathrm{p} R^T U = \mathrm{q} R^T \alpha U$ must hold when $k+1$ chooses A with probability p after $h \circ A$ and chooses B with probability q after $h \circ B$. Then, $q \in [1/\alpha R^T,\ 1]$ and $p \in [0,\ \alpha q - (1/R^T)]$ must hold. The other results are shown analogously. □

7. When $R^T > 1/\alpha$, then for all subgame perfect equilibria, players k and $k+1$ choose one of the following actions after $h \in H(B)$.

7.I. When k chooses B with probability $r \in (0,\ 1)$ after h, then $k+1$ chooses B with probability $q \in [0,\ \min\{1,\ (1/\alpha) - (1/R^T)\}]$ after $h \circ B$ and chooses A with probability $\alpha q + (\alpha/R^T)$ after $h \circ A$.

7.II. When k chooses B with probability one after h, then $k+1$ chooses B with probability $q \in [0,\ 1]$ after $h \circ B$ and chooses A with probability $p \in [0,\ \min\{1,\ \alpha \mathrm{q} + (\alpha/R^T)\}]$ after $h \circ A$.

7.III. When k chooses A with probability one after h, then $k+1$ chooses B with probability $q \in [0,\ \min\{1,\ (1/\alpha) - (1/R^T)\}]$ after $h \circ B$ and chooses A with probability $p \in [\alpha q + (\alpha/R^T),\ 1]$ after $h \circ A$.

Proof of 7.I. Because k is indifferent between B and A, $\alpha U + q R^T \alpha U = p R^T U$ must hold when $k+1$ chooses A with probability p after $h \circ A$ and chooses B with probability q after $h \circ B$. Then, $q \in [0,\ \min\{1,\ (1/\alpha) - (1/R^T)\}]$ and $p = \alpha q + (\alpha/R^T)$ must hold. The other results are shown analogously. □

It is especially shown that the following inefficient result holds.
8. (The inefficiency of equilibria): $\forall \epsilon > 0$, $\forall \alpha \in (0,\ 1)$, $\forall R > 1$, $\exists T(0)$, $\forall T \geq T(0)$, there exists a subgame perfect equilibrium where the average payoff of player $k \geq 1$ after any history is less than ϵ.

Proof. Take ϵ, α, and $R>1$. When T is sufficiently large, $R^T>1/\alpha$ and $1/(1+R^T)<\epsilon$ hold.
Consider the following action profiles.
Player 0 chooses A.
Player 1 chooses the same action as that of player 0.
For some $h \in H$, player $k \geq 2$ chooses B with probability $1/\alpha R^T$ after $h \circ A \circ B$, chooses B with probability one after $h \circ A \circ A$, chooses A with probability one after $h \circ B \circ B$, and chooses A with probability α/R^T after $h \circ B \circ A$.
Under these profiles, player $k \geq 1$ is indifferent between A and B, after any history.
The discounted payoffs of player $k \geq 1$ after $h \in H(A)$ and $h \in \mathrm{H}(B)$ are U and αU, and his average payoffs are $1/(1+R^T)<\epsilon$ and $\alpha/(1+R^T)<\epsilon$, respectively. □

Chapter 3

Rational Cooperation in the Prisoner's Dilemma with Finite Alternating Repetition

3.1 Introduction

In repeated games, players can obtain payoffs that are not obtainable in one-shot games. Under the condition of full dimensionality, Fudenberg and Maskin [12] showed that any feasible and individually rational payoffs of the stage game are attainable as subgame perfect equilibrium payoffs in repeated games. Many studies, including theirs, assume that the number of repetitions is infinite, which means the game does not end for all players. When there is an end to the game, efficient outcomes are sometimes not included in the set of subgame perfect equilibrium payoffs, because players always choose myopic behavior at the end of their lives. One reasonable way to avoid such myopic actions in finitely repeated games is to make a player afraid that his opponent may have to imitate his previous action. That is, with a certain probability, his opponent cooperates if he did so in the previous period, and punishes him if not. Such a strategy profile is called a Tit-for-Tat (TFT) strategy. Kreps et al. [17] considered the symmetric two-person prisoner's dilemma with simultaneous moves, where one player must play TFT, and showed that all sequential equilibria are almost efficient when the number of repetition is sufficiently large.

Motivated by these results, this chapter considers the symmetric two-person prisoner's dilemma with no discount and with finite alternating repetition. In the two-person prisoner's dilemma, two players, ROW and COL, choose either cooperation (C) or defection (D), where D strongly dominates C, and the payoff from cooperation is higher than that from defection by both players. In our model with alternating repetition, ROW can revise his action only in odd periods, whereas COL can only do so in even periods. We assume that ROW may have to play either the Grim-Trigger (GT) strategy, mentioned later, or a TFT strategy. At the beginning of the game, nature stochastically decides whether ROW is rational and able to choose his action freely, or has to play these strategy profiles. Only ROW observes his type directly, while the probability of his type is common knowledge. We examine whether and how players can attain efficient payoffs in this model when they repeat the game for a sufficiently long time. Despite the negative results in the literature, showing that inefficient outcomes that never arise in simultaneous move games sometimes cannot be excluded from the set of equilibrium payoffs in asynchronous move games, we show that *all* sequential equilibria still remain almost efficient, even though the structure of the repetition in our model is asynchronous.

When players' rationality is common knowledge in the finitely repeated prisoner's dilemma, they apply backward induction and always choose D in the last stage. Because players rationally expect this, they always choose D, for any history. As a result, only the inefficient outcome realizes, however large the number of repetitions is. This result may not reflect people's behavior in real world. For example, a player may commit to play the TFT strategy, and mimic what his opponent chose in the previous stage. Axelrod [17] compared the strategy profiles in the prisoner's dilemma, and showed that the TFT strategy has the highest payoff among the profiles. The study of Kreps et al. [17] is motivated by this result. One of the other representative strategy profiles that can make payoffs in finitely repeated games efficient is the Grim-Trigger (GT) strategy. According to this strategy, the player cooperates at first, and then continues to cooperate until his opponent deviates. Once the opponent deviates, the player continues to punish him throughout his life. Fudenberg and Maskin [12] studied this strategy in general two-person games with finite repetition, and proved the folk theorem by assuming that both players have to play the GT strategy, and that the number of repetitions is sufficiently large.

The proof of our result is based on that of Kreps et al. [17]. First, it is optimal for COL to always choose C, except in the last stage if ROW always plays either the GT or the TFT strategy. Second, if rational ROW chooses D when COL chooses C, ROW's rationality becomes common knowledge, and players continue to choose

D thereafter, yielding low continuation payoffs. However, if rational ROW chooses C when COL does so as well, then COL cannot tell the type of ROW and players can get higher continuation payoffs by cooperating. Therefore, it is optimal for COL to cooperate, except last few stages, if rational ROW also cooperates. Because ROW rationally expects, he also cooperates, except in the last few stages, if COL does so, which guarantees the existence of cooperative equilibria with efficient payoffs.

Our model of alternating repetition is a class of asynchronously repeated games. How the set of subgame perfect equilibria differs between synchronously and asynchronously repeated games has been researched for a long time. Some studies show that in some classes of asynchronously repeated games, players can avoid inefficient equilibria, which they cannot do in simultaneously repeated games. Lagunoff and Matsui [18] studied pure coordination games with alternating repetition, and showed that only the most efficient payoff is obtainable as a subgame perfect equilibrium payoff when players are sufficiently patient. Dutta [6] showed that in finitely repeated games of common interest, only the Pareto-efficient payoff arises in subgame perfect equilibria. However, other studies have shown that the infimum of equilibrium payoffs in some classes of repeated games with asynchronous repetition is lower than that with synchronous repetition. Takahashi and Wen [25] considered a specific three-person asynchronously repeated game, and showed that the minimum of average subgame perfect equilibrium payoffs for a player is less than his minimax value in the stage game. Wen [26] and Yoon [29] defined the effective minimax value of asynchronously repeated games that consider the structure of the repetition, and proved the folk theorem in which the infimum of individually rational payoffs is equal to this value, which is possibly less than the standard minimax value. The findings of these three studies imply that inefficient outcomes that do not occur in simultaneous move games sometimes cannot be excluded from the set of equilibrium payoffs when the repetition is asynchronous. Our result is closer to those of Lagunoff and Matsui [18] and Dutta [6], rather than those of Takahashi and Wen [25], Wen [26], and Yoon [29], in that inefficient outcomes do not appear in the equilibrium of our model.

In the remainder of this chapter, the abovementioned models are defined and the efficient results are proved.

3.2 Models and Efficient Results

3.2.1 Model 1: Grim-Trigger Strategy

The stage game considered in this chapter is the two-person prisoner's dilemma, where two players, ROW and COL, choose either C (Cooperate) or D (Defect). If both players choose C, their payoff is 1, whereas the payoff is 0 if they both choose D. If one player chooses D and the other chooses C, their payoffs are a and b, respectively, where $a > 1$, $b < 0$, and $a + b < 2$. The finitely repeated game with alternating repetition is defined as follows. COL can revise his action only in even periods, including period 0, whereas ROW can only do so in odd periods. The number of repetitions from period 1 to the last period is an even number N. It is assumed that players do not discount the future and that COL gets no payoff at period 0. At the beginning of the game, nature assigns probability $q_{GT} \in (0,\ 1)$ to the event that ROW has to play the Grim-Trigger (GT) strategy, and assigns probability $1 - q_{GT}$ to ROW being rational and able to choose his action without any restriction. Only ROW directly observes nature's decision, whereas the probability q_{GT} on ROW's type is common knowledge. We define *GT-player* as the player who continues to choose C until his opponent chooses D, and then continues to choose D (say, $all - D$) once his opponent has chosen D. The remainder of this subsection proves the following efficiency.

Efficiency (GT). In every sequential equilibrium of this game, each player's total payoff is at least $N + b - 3 - \frac{2}{q_{GT}}(a - b - 1)$, where the per-period average converges to 1 as $N \to \infty$.

The proof resembles that of Kreps et al. [17].

Step 1. If ROW either chooses D, even though COL has been choosing C since the beginning of the game, or chooses C after COL chooses D, both players choose D at every subsequent stage.
– In this case, the continuation game becomes the finitely repeated prisoner's dilemma, with alternating repetition by rational players. Players apply backward induction and play $all - D$ throughout subsequent stages.

Step 2. When COL chooses D, ROW chooses D at the same stage.
– If ROW chooses C at this stage, his rationality is revealed, and his continuation payoff is $2b$, from step 1. However, if he chooses D at this stage and plays $all - D$ thereafter, his continuation payoff is at least $0 > 2b$.

Step 3. After COL chooses D, ROW plays $all - D$.
– Suppose COL chooses D and then at a later stage, chooses C. If ROW chooses C at this stage, his continuation payoff is $1 + b$, from step 1. If he chooses D at this stage, his continuation payoff is at least a. Therefore, ROW continues to choose D once COL chooses D.

Step 4. After one player chooses D, players play $all - D$.
– By step 1, players play $all - D$ if ROW deviates first. If COL deviates first, ROW plays $all - D$, from step 3, so COL's best response is to also play $all - D$.

Step 5. The difference in continuation payoffs between players at any history is at most $a - b$.
– Players' one-shot payoffs differ only when one player chooses C and the other chooses D, which can happen only once throughout the game, by step 4.

Step 6. When there are n stages left and COL's assignment on GT-ROW at this stage is q, COL's continuation payoff is at least $1 + q(n - 2 + a) + (1 - q)b$ if players have been choosing C since the beginning of the game.
– If COL plays GT for $n - 1$ periods, and then chooses D at the last stage, his continuation payoff is $n - 1 + a$ against GT-ROW with probability q, whereas it is at least $1 + b$ against rational ROW with probability $1 - q$.

Step 7. When there are n stages left and COL's assignment on GT-ROW at the stage following $(C,\ C)$ is q, ROW's continuation payoff is at least $2 + q(n - 3 + a) + (1 - q)b - a + b$ if players have been choosing C since the beginning of the game.
– When ROW chooses C at this stage, he gets 1 immediately. Because he keeps choosing C at the next stage, his continuation payoff for the following $n - 1$ stages is at least $1 + q(n - 3 + a) + (1 - q)b - a + b$, by steps 5 and 6.

Step 8. When there are $n > \frac{2}{q}(a - b - 1) - a + b + 3$ stages left and COL's assignment on GT-ROW at the stage following $(C,\ C)$ is $q > 0$, ROW chooses C if the players have been doing so since the beginning of the game.
– If ROW chooses D at this stage, his continuation payoff is a, by step 4. On the other hand, his continuation payoff must be at least $2 + q(n - 3 + a) + (1 - q)b - a + b$, by step 7, which is strictly greater than a because of the definition of n. Therefore, ROW chooses C at this stage.

Step 9. When there are $n > \frac{2}{q_{GT}}(a - b - 1) - a + b + 3$ stages left, ROW chooses C if players have been doing so since the beginning of the game.
– This is because COL's belief on GT-ROW remains unchanged, and is equal to q_{GT} throughout these stages, by step 8.

Step 10. COL's total payoff is at least $N + a - 3 - \frac{2}{q_{GT}}(a - b - 1)$.
– If COL plays GT, his total payoff is at least this value, by step 9.

Step 11. ROW's total payoff is at least $N + b - 3 - \frac{2}{q_{GT}}(a - b - 1)$.
– By steps 5 and 10, ROW's total payoff is at least this value. □

3.2.2 Model 2: Tit-for-Tat Strategy

In this subsection, we assume that nature assigns probability q_{TFT} at the beginning of the game to ROW playing the Tit-For-Tat (TFT) strategy. We define *TFT-player* in our alternately repeated game as the player who mimics what his opponent *is choosing*. The following efficiency holds.

Efficiency (TFT). In every sequential equilibrium of this game, each player's total payoff is at least $N + b - 3 - \frac{2}{q_{TFT}}(a - b - 1)$.

The proof is similar to that in the previous subsection.

Step 1. If ROW chooses a different action to that of COL, players choose D at every subsequent stage.

Step 2. When COL chooses D, ROW chooses D at the same stage.

Step 3. The difference between the continuation payoffs of the players for any history is at most $a - b$ if ROW plays TFT.
– If ROW plays TFT, the players' payoffs differ only when COL does not make the same choice as ROW. If COL wants to choose D twice when ROW is choosing C, he must choose C when ROW chooses D, offsetting the difference in the payoffs from COL's first deviation.

Step 4. For any history, COL's continuation payoff if ROW plays TFT is at least that when ROW plays an equilibrium strategy.
– Suppose there are two stages left. When COL chooses D, ROW chooses D in both cases, by step 2, and COL's

continuation payoff does not differ. When COL chooses C, TFT-ROW also chooses C, while rational ROW chooses D. Next, suppose there are n stages left and that step 4 holds for continuation games with $n-1$ stages remaining. All we need to be concerned about is when ROW can move at this stage, because the proof when ROW cannot move can be shown analogously. When COL chooses D, ROW chooses D in both cases. When ROW plays TFT and COL chooses C, COL gets 1 at this stage and his continuation payoff for the subsequent $n-1$ stages is at least the same as when ROW plays TFT thereafter. When ROW plays an equilibrium strategy and COL chooses C, COL gets 1 at this stage and his continuation payoff for subsequent stages is at most that against TFT-ROW if ROW chooses C, whereas COL gets b in total if ROW chooses D at this stage.

Step 5. If there are n stages left and COL's assignment on TFT-ROW at this stage is q, COL's continuation payoff when choosing C is at least $1+q(n-2+a)+(1-q)b$.

Step 6. If there are n stages left and COL's assignment on TFT-ROW in the stage after $(C,\ C)$ is played is q, ROW's continuation payoff when COL chooses C is at least $2+q(n-3+a)+(1-q)b-a+b$.
– If ROW chooses C at this stage, he gets 1 immediately and keeps choosing C at the next stage. By the definition of an equilibrium strategy, ROW's continuation payoff for the subsequent $n-1$ stages when he plays an equilibrium strategy is at least the same as when he plays TFT. Then, we can see this observation, applying steps 3 to 5 in order.

Step 7. If there are $n > \frac{2}{q}(a-b-1)-a+b+3$ stages left and COL's assignment on TFT-ROW at the stage after $(C,\ C)$ is played is $q>0$, ROW chooses C when COL does the same.

Step 8. If there are $n > \frac{2}{q_{TFT}}(a-b-1)-a+b+3$ stages left, ROW chooses C when COL does the same.

Step 9. COL's total payoff is at least $N+a-3-\frac{2}{q_{TFT}}(a-b-1)$.

Step 10. ROW's total payoff is at least $N+b-3-\frac{2}{q_{TFT}}(a-b-1)$.
– If ROW plays TFT throughout the game, he can get at least this value, by steps 3 and 9. □

Chapter 4

Efficiency in the Repeated Prisoner's Dilemma with Private Monitoring when the Distribution of Signals Changes over Time

4.1 Introduction

Infinitely repeated games have been broadly applied to economic models to show that inefficient outcomes can be avoided through repetition. In order to punish myopic defections, however, players' common observations about their past play have been thought to be necessary. Many studies on repeated games, have assumed that the monitoring is either perfect or imperfect but public, and have chiefly studied (public) perfect equilibria. In the case of repeated games with imperfect public monitoring, Fudenberg and Maskin [10] proved the folk theorem.

Repeated games where players' actions are not observable to each other, but where each player receives private signals with noise, which depend on his opponents' previous actions, are natural settings for many economic models in the real world. This chapter considers the two-person repeated prisoner's dilemma with private monitoring, where the probability distribution of private signals changes over time, depending on calendar dates. In the two-person prisoner's dilemma, players choose either cooperation (C) or defection (D), where D strongly dominates C, and the payoff from cooperation is higher than that from defection. At the end of every period, each player receives a noisy private signal with a certain probability, depending on his opponent's current play, rather than on directly observing his opponent's actions. We prove that efficient belief-free sequential equilibria exist with sufficiently patient players and when the monitoring is almost perfect in every period.

In order to prove this result, we consider the belief-free strategy, which is a representative class of strategy profiles in repeated games with private monitoring. This approach focuses on a set of continuation strategy profiles possessing the property that for any possible belief of the opponent's continuation strategy in the set, any continuation strategy in the set is optimal for a player. Given such a construction, players become indifferent among all their continuation strategy profiles, which is independent of their private observations. As a result, their beliefs have nothing to do with verifying optimality. Belief-free sequential equilibria in repeated games with private monitoring was first introduced by Piccione [22]. Ely and Välimäki [8] considered the two-person repeated prisoner's dilemma with private, almost perfect monitoring, and proved the folk theorem. Ely et al. [7] applied this result to general two-person games with private monitoring, and characterized the set of belief-free equilibria when the noise vanishes. Hörner and Olszewski [14] generalized this result to n-person games and proved the folk theorem under almost perfect monitoring.

In our model, we construct a Markov strategy, as in Ely and Välimäki [8], under an invariant monitoring structure, where the distribution of signals is fixed throughout the repetitions. The Markov strategy in our model is a class of belief-free strategy profiles that depends only on each player's play and observation in the previous period, and the calendar date. It is shown that when a player's opponent adopts a Markov strategy, he is indifferent between C and D in every period and, therefore, cannot change his continuation payoff by himself. When players are sufficiently patient and the monitoring technology is almost perfect every period, such a strategy exists for the payoff vector between the payoff of (C, C) and (D, D), which ensures that efficient payoffs are attainable in this game.

The model considered in this chapter with variant monitoring structure can bee seen as a repeated stochastic game, where the circumstances of the stage game in each period change over time. In repeated stochastic games with public monitoring, where the state of the game in the next period changes according to the probability distribution, which depends on realized actions, the public signal, and the state in the current period, Hörner et al. [15] characterized the set of public perfect equilibria and proved the folk theorem. Aiba [1] studied stochastic games with private monitoring and proved the folk theorem.

In the literature on stochastic games, it is usually assumed that the number of states is finite and that the transition of states in each period is determined endogenously. That is, one of finitely many states emerges every period, depending only on the realized variables in the previous period. Our study, on the other hand, focuses on the case where the distribution of private signals changes exogenously, depending absolutely on calendar dates. We also admit the case where the distribution of signals in any period is different from any other period.

We also apply this result to alternately repeated games. One of the primary works on asynchronously repeated games is that of Maskin and Tirole [20] [21], who modeled an imperfectly competitive market as a repeated game between duopolists with alternating moves, which is a canonical form of asynchronous repetition. They analyzed Markov perfect equilibria in quantity or price competition, which are played alternately. Following this work, the asynchronous repetition of various kinds of stage games has been studied. Under nonequivalent utility, finite periods of inaction conditions, and the assumptions that past mixed actions are perfectly observable and players who do not move in a period must play the same mixed actions as in the previous period, Yoon [28] obtained the folk theorem in asynchronously repeated games.

Based on these results, we consider the alternately repeated prisoner's dilemma with private monitoring. In this game, player 1 can change his action only in odd periods, while player 2 can only do so in even periods. We are able to construct efficient belief-free Markov strategy profiles that resemble that in the simultaneously repeated game. In contrast to Yoon [28], we do not use public randomizations.

The rest of this chapter is organized as follows. Section 2 defines the model of simultaneous repetitions and proves the efficiency. Section 3 considers the case of alternating repetition.

4.2 The Model and Efficiency

We consider the stage game of the two-person prisoner's dilemma, where two players choose either C (Cooperate) or D (Defect). Denote player $i \in \{1,\ 2\}$'s one-shot expected payoff as $u^i_{a_i a_j}$ when the realized action of i and $j \neq i$ is $(a_i,\ a_j) \in \{C,\ D\}^2$. We assume $u^i_{DC} > u^i_{CC} > u^i_{DD} > u^i_{CD}$ and $u^i_{DC} + u^i_{CD} < 2u^i_{CC}$ for every i. At each period $t \geq 1$, the game with private monitoring is repeated, where each player i receives a private signal ω_i of either c (good) or d (bad). When players choose $(a_1,\ a_2)$ at period t, the pair of private signals $(\omega_1,\ \omega_2) \in \{c,\ d\}^2$ is realized with probability $m^t_{a_1 a_2}(\omega_1,\ \omega_2) \in (0,\ 1)$. We use a discount factor $\delta \in (0,\ 1)$ to discount players' payoffs at period t by $(1-\delta)\delta^{t-1}$. Player i's private history at the beginning of period t, say $h^{i,t}$, is a $2(t-1)$-length sequence of what he played and observed in previous periods. Define i's marginal probability distribution of the good signal at period t when $(a_i,\ a_j)$ is taken, as follows.

$$m^{i,t}_{a_i a_j} = \textstyle\sum_{\omega_j \in \{c,d\}} m^t_{a_1 a_2}(c,\ \omega_j).$$

For $\epsilon > 0$, the monitoring structure is called ϵ-perfect when, for all t, $m^{i,t}_{a_i C} \geq 1-\epsilon$ and $m^{i,t}_{a_i D} \leq \epsilon$ hold for every i and a_i. In this chapter, we consider the strategy profile in which each player i chooses C at period 1 and then plays a Markov strategy $\pi^{i,t}_{a_i \omega_i} \in (0,\ 1)$, with which probability he chooses C at period $t \geq 2$ when he chose a_i and observed ω_i at period $t-1$. Given such a strategy profile, let $\mu^i(h^{i,t})[a_j]$ be his consistent belief, calculated by Bayes' rule, with which probability j chooses a_j after $h^{i,t}$. Player i's continuation payoff after history $h^{i,t}$, where he chose a_i and observed ω_i in the previous period, say $w^i(h^{i,t})$, is written as follows:

$$\begin{gathered} w^i(h^{i,t}) = \textstyle\sum_{a_j} \mu^i(h^{i,t})[a_j](\pi^{i,t}_{a_i\omega_i} w^{i,t}_{Ca_j} + (1-\pi^{i,t}_{a_i\omega_i}) w^{i,t}_{Da_j}), \\ \text{where } w^{i,t}_{a_i a_j} = (1-\delta)u^i_{a_i a_j} + \delta \textstyle\sum_{(\omega_1,\omega_2)} m^t_{a_1 a_2}(\omega_1,\ \omega_2) \\ (\pi^{j,t+1}_{a_j\omega_j}(\pi^{i,t+1}_{a_i\omega_i} w^{i,t+1}_{CC} + (1-\pi^{i,t+1}_{a_i\omega_i}) w^{i,t+1}_{DC}) + (1-\pi^{j,t+1}_{a_j\omega_j})(\pi^{i,t+1}_{a_i\omega_i} w^{i,t+1}_{CD} + (1-\pi^{i,t+1}_{a_i\omega_i}) w^{i,t+1}_{DD})). \end{gathered}$$

The following efficient result holds.

Efficiency. $\forall (v^1_C,\ v^2_C) \in (u^1_{DD},\ u^1_{CC}) \times (u^2_{DD},\ u^2_{CC})$ and $\exists \delta_0 \in (0,\ 1)\ \forall \delta \geq \delta_0\ \exists \epsilon > 0$, and for all ϵ-perfect monitoring structures, there exists a sequential equilibrium with a payoff vector $(v^1_C,\ v^2_C)$.

Proof. For $(a_i,\ a_j) \in \{C,\ D\}^2$, consider the following Markov strategy for $j \in \{1,\ 2\}$:

$$\begin{aligned} \pi^{j,t+1}_{a_j c} &= \frac{(1-m^{j,t}_{a_j D})(v^i_{a_j} - \delta v^i_D - (1-\delta)u^i_{Ca_j}) - (1-m^{j,t}_{a_j C})(v^i_{a_j} - \delta v^i_D - (1-\delta)u^i_{a_i D})}{\delta(m^{j,t}_{a_j C} - m^{j,t}_{a_j D})(v^i_C - v^i_D)}, \\ \pi^{j,t+1}_{a_j d} &= \frac{m^{j,t}_{a_j C}(v^i_{a_j} - \delta v^i_D - (1-\delta)u^i_{a_i D}) - m^{j,t}_{a_j D}(v^i_{a_j} - \delta v^i_D - (1-\delta)u^i_{Ca_j})}{\delta(m^{j,t}_{a_j C} - m^{j,t}_{a_j D})(v^i_C - v^i_D)}. \end{aligned}$$

When players are sufficiently patient and the monitoring at every period is almost perfect, these values lie on $(0,\ 1)$ because they are positive and their denominator is greater than their numerator. We can rewrite these variables as follows:

$$
\begin{aligned}
v^i_{a_j} &= (1-\delta)u^i_{a_i a_j} \\
&+(m^{j,t}_{a_j a_i}\pi^{j,t+1}_{a_j c} + (1-m^{j,t}_{a_j a_i})\pi^{j,t+1}_{a_j d})\delta v^i_C + (m^{j,t}_{a_j a_i}(1-\pi^{j,t+1}_{a_j c}) + (1-m^{j,t}_{a_j a_i})(1-\pi^{j,t+1}_{a_j d}))\delta v^i_D.
\end{aligned}
$$

Then we can see the following relationships hold for a_j and t.

$$w^{i,t}_{Ca_j} = w^{i,t}_{Da_j} = v^i_{a_j}.$$

Therefore, the following equation holds for each player $i \neq j$:

$$w^i(h^{i,t}) = \mu^i(h^{i,t})[C]v^i_C + (1-\mu^i(h^{i,t})[C])v^i_D.$$

Given j's strategy, this equation implies that whatever i chooses at period t, his continuation payoff does not change. Thus, each player i's strategy, as defined above, is a best response to j's choice, forming an equilibrium strategy. Their total expected payoffs are $(w^1,\ w^2) = (v^1_C,\ v^2_C)$ because they play C with probability 1 at period 1. □

4.3 Alternating Repetition

In this section, we consider the game with alternating repetition. Player 1 can revise his action only in odd periods, while player 2 can only do so in even periods and in period 1. We assume that at period 1, player 2 can only choose his action and gets no payoff. The monitoring structure and sequence of signal distributions from period 1 is defined in the same way as in the previous section. Player i's continuation payoff after history $h^{i,t}$ when he can move, except in period 1 for player 2, is as follows:

$$
\begin{aligned}
w^i(h^{i,t}) &= \textstyle\sum_{a_j} \mu^i(h^{i,t})[a_j](\pi^{i,t}_{a_i\omega_i} w^{i,t}_{Ca_j} + (1-\pi^{i,t}_{a_i\omega_i})w^{i,t}_{Da_j}), \\
\text{where } w^{i,t}_{a_i a_j} &= (1-\delta)u^i_{a_i a_j} \\
&+\delta(m^{j,t}_{a_j a_i}\pi^{j,t+1}_{a_j c} + (1-m^{j,t}_{a_j a_i})\pi^{j,t+1}_{a_j d})(1-\delta)u^i_{a_i C} \\
&+\delta(m^{j,t}_{a_j a_i}(1-\pi^{j,t+1}_{a_j c}) + (1-m^{j,t}_{a_j a_i})(1-\pi^{j,t+1}_{a_j d}))(1-\delta)u^i_{a_i D} \\
&+\delta^2(m^{j,t}_{a_j a_i}\pi^{j,t+1}_{a_j c} + (1-m^{j,t}_{a_j a_i})\pi^{j,t+1}_{a_j d})(m^{i,t+1}_{a_i C}\pi^{i,t+2}_{a_i c} + (1-m^{i,t+1}_{a_i C})\pi^{i,t+2}_{a_i d})w^{i,t+2}_{CC} \\
&+\delta^2(m^{j,t}_{a_j a_i}\pi^{j,t+1}_{a_j c} + (1-m^{j,t}_{a_j a_i})\pi^{j,t+1}_{a_j d})(m^{i,t+1}_{a_i C}(1-\pi^{i,t+2}_{a_i c}) + (1-m^{i,t+1}_{a_i C})(1-\pi^{i,t+2}_{a_i d}))w^{i,t+2}_{DC} \\
&+\delta^2(m^{j,t}_{a_j a_i}(1-\pi^{j,t+1}_{a_j c}) + (1-m^{j,t}_{a_j a_i})(1-\pi^{j,t+1}_{a_j d}))(m^{i,t+1}_{a_i D}\pi^{i,t+2}_{a_i c} + (1-m^{i,t+1}_{a_i D})\pi^{i,t+2}_{a_i d})w^{i,t+2}_{CD} \\
&+\delta^2(m^{j,t}_{a_j a_i}(1-\pi^{j,t+1}_{a_j c}) + (1-m^{j,t}_{a_j a_i})(1-\pi^{j,t+1}_{a_j d}))(m^{i,t+1}_{a_i D}(1-\pi^{i,t+2}_{a_i c}) + (1-m^{i,t+1}_{a_i D})(1-\pi^{i,t+2}_{a_i d}))w^{i,t+2}_{DD}.
\end{aligned}
$$

The following result holds.

Efficiency with alternating repetition. $\forall(v^1_C,\ v^2_C) \in (u^1_{DD},\ u^1_{CC}) \times (u^2_{DD},\ u^2_{CC})$ and $\exists\delta_0 \in (0,\ 1)$ $\forall\delta \geq \delta_0$ $\exists\epsilon > 0$, and for all ϵ-perfect monitoring structures, there exists a sequential equilibrium with a payoff vector $(v^1_C,\ v^2_C)$ in this game with alternating repetition.

Proof. Consider j's Markov strategy:

$$
\begin{aligned}
\pi^{j,t+1}_{a_j c} &= \frac{(1-m^{j,t}_{a_j D})(v^i_C-v^i_D+(1-\delta)(u^i_{DC}-u^i_{DD}))(v^i_{a_j}-\delta v^i_D-(1-\delta)(u^i_{Ca_j}+\delta u^i_{CD}))}{\delta(m^{j,t}_{a_j C}-m^{j,t}_{a_j D})(v^i_C-v^i_D+(1-\delta)(u^i_{CC}-u^i_{CD}))(v^i_C-v^i_D+(1-\delta)(u^i_{DC}-u^i_{DD}))} \\
&-\frac{(1-m^{j,t}_{a_j C})(v^i_C-v^i_D+(1-\delta)(u^i_{CC}-u^i_{CD}))(v^i_{a_j}-\delta v^i_D-(1-\delta)(u^i_{a_i D}+\delta u^i_{DD}))}{\delta(m^{j,t}_{a_j C}-m^{j,t}_{a_j D})(v^i_C-v^i_D+(1-\delta)(u^i_{CC}-u^i_{CD}))(v^i_C-v^i_D+(1-\delta)(u^i_{DC}-u^i_{DD}))}, \\
\pi^{j,t+1}_{a_j d} &= \frac{m^{j,t}_{a_j C}(v^i_C-v^i_D+(1-\delta)(u^i_{CC}-u^i_{CD}))(v^i_{a_j}-\delta v^i_D-(1-\delta)(u^i_{a_i D}+\delta u^i_{DD}))}{\delta(m^{j,t}_{a_j C}-m^{j,t}_{a_j D})(v^i_C-v^i_D+(1-\delta)(u^i_{CC}-u^i_{CD}))(v^i_C-v^i_D+(1-\delta)(u^i_{DC}-u^i_{DD}))} \\
&-\frac{m^{j,t}_{a_j D}(v^i_C-v^i_D+(1-\delta)(u^i_{DC}-u^i_{DD}))(v^i_{a_j}-\delta v^i_D-(1-\delta)(u^i_{Ca_j}+\delta u^i_{CD}))}{\delta(m^{j,t}_{a_j C}-m^{j,t}_{a_j D})(v^i_C-v^i_D+(1-\delta)(u^i_{CC}-u^i_{CD}))(v^i_C-v^i_D+(1-\delta)(u^i_{DC}-u^i_{DD}))}.
\end{aligned}
$$

When players are sufficiently patient and the monitoring is almost perfect at every period, these values lie on $(0,\ 1)$. Rewrite these variables as follows:

$$
\begin{aligned}
v^i_{a_j} &= (1-\delta)u^i_{a_i a_j} \\
&+\delta(m^{j,t}_{a_j a_i}\pi^{j,t+1}_{a_j c} + (1-m^{j,t}_{a_j a_i})\pi^{j,t+1}_{a_j d})((1-\delta)u^i_{a_i C} + \delta v^i_C) \\
&+\delta(m^{j,t}_{a_j a_i}(1-\pi^{j,t+1}_{a_j c}) + (1-m^{j,t}_{a_j a_i})(1-\pi^{j,t+1}_{a_j d}))((1-\delta)u^i_{a_i D} + \delta v^i_D).
\end{aligned}
$$

Then, we find that the following relationships hold:

$$w^{i,t}_{Ca_j} = w^{i,t}_{Da_j} = v^i_{a_j}.$$

Therefore, the following equation holds:

$$w^i(h^{i,t}) = \mu^i(h^{i,t})[C]v^i_C + (1-\mu^i(h^{i,t})[C])v^i_D.$$

Their total expected payoffs are $(v^1_C,\ v^2_C)$ when they play C with probability 1 at period 1. □

Bibliography

[1] Aiba, K. (2014), "A Folk Theorem for Stochastic Games with Private Almost-Perfect Monitoring," *Games and Economic Behavior* 86, 58–66.

[2] Axelrod, R. (1984), "*The Evolution of Cooperation*," Basic Books, New York.

[3] Bailey, K., and R. Beard, "N-Organization Overlapping Generations Games with Memoryless Agents," *Economic Theory*.

[4] Benoit, J., and V. Krishna (1985), "Finitely Repeated Games," *Econometrica* 53, 905–922.

[5] Bhaskar, V. (1998), "Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems," *Review of Economic Studies* 65, 135–149.

[6] Dutta, P. (2012), "Coordination Need not be a Problem," *Games and Economic Behavior* 76, 519–534.

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

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

[9] Fudenberg, D., D. Kreps, and E. Maskin (1990), "Repeated Games with Long-Run and Short-Run Players," *Review of Economic Studies* 57, 555–573.

[10] Fudenberg, D., and D. Levine (1994), "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory* 62, 103–135.

[11] Fudenberg, D., D. Levine, and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica* 62, 997–1040.

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

[13] Gossner, O. (1996), "Overlapping Generations Games with Mixed Strategies," *Mathematics of Operations Research* 21, 477–486.

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

[15] Hörner, J., S. Takahashi, T. Sugaya, and N. Vieille (2011), "Recursive Methods in Discounted Stochastic Games:
An Algorithm for $\delta \to 1$ and a Folk Theorem," *Econometrica* 79, 1277–1318.

[16] Kandori, M. (1992), "Repeated Games Played by Overlapping Generations of Players," *Review of Economic Studies* 59, 81–92.

[17] Kreps, D., P. Milgrom, J. Roberts and R. Wilson (1982), "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory* 27, 245–252.

[18] Lagunoff, R., and A. Matsui (1997), "Asynchronous Choice in Repeated Coordination Games," *Econometrica* 65, 1467–1477.

[19] Lagunoff, R., and A. Matsui (2004), "Organizations and Overlapping Generations Games: Memory, Communication, and Altruism," *Review of Economic Design* 8, 383–411.

[20] Maskin, E., and J. Tirole (1988), "A Theory of Dynamic Oligopoly I: Overview and Quantity Competition with Large Fixed Costs," *Econometrica* 56, 549–570.

[21] Maskin, E., and J. Tirole (1988), "A Theory of Dynamic Oligopoly II: Price Competition, Kinked Demand

Curves and Fixed Costs," *Econometrica* 56, 571–600.

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

[23] Salant, D. (1991), "A Repeated Game with Finitely Lived Overlapping Generations of Players," *Games and Economic Behavior* 3, 244–259.

[24] Smith, L. (1992), "Folk Theorems in Overlapping Generations Games," *Games and Economic Behavior* 4, 426–449.

[25] Takahashi, S., and Q. Wen (2003), "On Asynchronously Repeated Games," *Economics Letters* 79, 239–245.

[26] Wen, Q. (2002), "A Folk Theorem for Repeated Sequential Games," *Review of Economic Studies* 69, 493–512.

[27] Yoon, K. (2001), "An Anti-Folk Theorem in Overlapping Generations Games with Limited Observability," *Review of Economic Dynamics* 4, 736–745.

[28] Yoon, K. (2003), "A Folk Theorem for Asynchronously Repeated Games," *Econometrica* 69, 191–200.

[29] Yoon, K. (2003), "The Effective Minimax Value of Asynchronously Repeated Games," *International Journal of Game Theory* 32, 431–442.