The Pennsylvania State University
The Graduate School
College of the Liberal Arts

ESSAYS ON DYNAMIC GAMES AND FORWARD INDUCTION

A Dissertation in
Economics
by
Shigeki Isogai

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2017

The dissertation of Shigeki Isogai was reviewed and approved* by the following:

Edward Green
Professor Emeritus of Economics
Dissertation Co-Advisor, Co-Chair of Committee

Robert C. Marshall
Distinguished Professor of Economics
Dissertation Co-Advisor, Co-Chair of Committee

Vijay Krishna
Distinguished Professor of Economics

Lisa Posey
Associate Professor of Risk Management

Barry W. Ickes
Professor of Economics
Head of the Department of Economics

*Signatures are on file in the Graduate School.

Abstract

In this essay, I study how forward-induction reasoning affect plausibility/stability of agreements in which players in a dynamic interaction enforces cooperation with the threat of mutually destructive punishment. While the traditional theory using equilibrium concept shows that such strategy profile is self-enforcing, under a modification of the model, such strategy profile fails to be consistent with players' rationality.

In the first chapter I provide the simplest setting under which this non-rationalizability result of deterrence can be shown. The game is a two-player three-stage game: in the first stage, the players choose whether to enter the strategic interaction by paying some cost; in the second stage, the players play a prisoners' dilemma game; and in the third stage, the players play a 2×2 coordination game. Each move is simultaneous and the players' past actions are perfectly monitored. While there exists a subgame-perfect equilibrium in which players can cooperate with the threat of punishment provided the punishment is strong enough, I show that the strategy profile does not consists of rationalizable strategies under a certain parameter values. This occurs because choosing to enter, unilaterally defect, and then punish the opponent is strictly dominated by a mixture of the two strategies "do not enter" and "enter, defect, but do not punish." This result shows that a simple modification of the game and forward-induction consideration encoded in rationalizability might cast doubt on the idea of deterring defection by the threat of mutual punishment.

The other two chapters study to what extent the result in the fist chapter does or does not apply in different settings. The second chapter considers the infinite-horizon extension of the model in the first chapter. In the first period (denoted as period 0), the players choose whether to enter the game. After the players choose to enter, the continuation game is the infinite repetition of the stage game which consists of two phases: in the first phase players play prisoners' dilemma game, after which players simultaneously choose to continue the game, exit from the game without punishing the opponent, or punish the opponent and

exit from the game. I show that with a similar condition as in the result in the first chapter, strategy which entails defection and punishment in the first stage is not rationalizable. Moreover, since the exit-without-punishment option works as an outside option in later stages of the game, we also obtain a result which provides conditions under which punishment after defection is excluded by rationalizability.

The third chapter extends the model in the first chapter toward an incomplete-information model in that it considers a model of random number of players, who are sequentially matched and play the game as in the first chapter. I assume that while the past actions in the stage games are not observable, occurrences of punishment is publicly observable to all the players (the typical example is the formation of cartels and the occurrence of leniency applications). I explore how this observable punishment works as a signaling device and how this model gives rise to a rationalizable use of punishment. I first show that a simple repetition of games does not give rise to a rationalizable punishment because of the assumption that the players cannot distinguish the non-occurrence of deviation and failure to punishment. I then discuss possible modifications to recover the punishment being an equilibrium action; i.e., that a small perturbation in payoffs can recover the possibility of punishment.

Table of Contents

List of Tables

List of Symbols

$\mathcal{H}$	The set of non-terminal histories.
$\mathcal{Z}$	The set of terminal histories.
Σ	The set of (pure) strategies
$\Sigma(h)$	The set of strategies consistent with history h
$\hat{U}$	Extensive-form payoff function
U	Strategic-form payoff function
$\mathcal{C}$	The set of conditional probability systems.
Δ	A restriction to players' possible beliefs.
$\Sigma(\Delta, n)$	The set of n-th round Δ-rationalizable strategies.
$\Sigma(\Delta)$	The set of Δ-rationalizable strategies.

Acknowledgments

I cannot thank Ed Green more for his help during my PhD study. His advice has provided me invaluable insights on how I should work with economic researches and how I should see the society as a social scientist. Without his advice I could not have completed my PhD study.

Bob Marshall, as my co-advisor, has been a great advisor to me especially in terms of applied studies. He has given me the sense in which how applied economic models should be constructed and how I should convey the relationship between the economic models and reality.

I also would like to express my indebtedness to Vijay Krishna and Kalyan Chatterjee, who have given me lots of important advice. It is my greatest pleasure to be able to study in such a great environment in microeconomic studies as Penn State.

I also thank many current and former PhD students at Penn State, to name a few, Yu Awaya, Tetsuya Hoshino, Gunhaeng Lee, Ryo Makioka, Wataru Nozawa, Mengli Sha, Chaohai Shen, Yuta Watabe (in alphabetical order); they have been good advisors to discuss research with and great friends to chat with. In particular, my coauthor Chaohai Shen has been a great friend and a researcher. I have benefited from discussion with him not only in our joint work but also in the papers in this dissertation.

John Riew has helped me come to Penn State and has been a great mentor for Asian PhD students including me.

Krista Winkelblech has helped me completing paperwork related to the graduate study. Without her help, my life here would be much tougher and more complicated.

Atsushi Kajii, the advisor in my Master's study, helped me continue my PhD study at Penn State. Takekazu Iwamoto, the advisor in my Bachelor's study, has made me interested in this difficult but deep research field. I owe them so much that I am here at Penn State thanks to their help.

Last, but not least, I am grateful to my parents, my sister, and my grandparents for being so patient to me.

I could complete the PhD program thanks to those who are listed above (and people whom I could not list here).

All remaining errors are my own.

Chapter 1
Credibility of Deterrence and Extensive-Form Rationalizability

1.1 Introduction

The idea of deterrence is prominent in international relations (mutually assured destruction, or MAD), industrial organization (self-enforcing collusion), and other areas of social science to which game theory has been applied. For example, two superpowers might be going to be engaged in warfare when their massive destructive weapons might bring significant harm to people in both countries. Each of them might be tempted to directly attack each other, but the threat of more massive full-fledged war might deter the direct attack to avoid catastrophe. Also in self-enforcing cartels, it is a well-known result that a temptation to secret price cutting is deterred by the possibility of mutual punishment, called "trigger-strategy."

In this article, I consider a simple finite extensive-form model of deterrence that can be interpreted, for example, as a model of MAD or as a reduced-form model of self-enforcing collusion. A crucial feature of this model is that each of the players initially makes an entry decision, that is, a decision to bear a cost in order to play the game. In international relations, an offensive-weapon system that can be used either for deterrence or for aggression may be more costly to build than a defensive fortification. In industrial organization, a firm may have to build a costly factory before exercising an option to sell a good either competitively or collusively; or, the cost of participation might be an expected loss from detection by the antitrust authority.

I apply the insight of prior game theory that an entry cost can have implications, via "forward induction" considerations, for what constitutes plausible behavior in the post-entry game. Specifically, I show that, for some parameter values, deterrence is not an extensive-form rationalizable strategy (Pearce, 1984) in a game with costly entry.

More precisely, I consider a model where a participation stage in which two players decide whether or not to enter a strategic relationship precedes a Prisoners' Dilemma (PD) game followed by a symmetric coordination game with Pareto-ranked equilibria. A strategy profile that supports cooperation in PD game by the threat of the play in Pareto-dominated outcome, which we call "punishment" here, is a subgame-perfect equilibrium (SPE), given that the deterrence is sufficiently strong. I show that, for some parameter values, this SPE is excluded by extensive-form rationalizability.

The intuition of the result is extremely simple: if one player, let us say "he," deviates from the cooperative action in the PD game, the opponent, "she," consider whether or not to punish him after observing the deviation. Under a certain circumstance, she realizes that he should not have participated in the strategic relationship had he intended to trigger a punishment after his deviation in the PD game because that is an inferior strategy for him. Thus, as long as she believes that he has rationally deviated in the PD game and not intended to trigger a punishment, she will be better off by not punishing him, either.

The model and the result is simple, but they are strong enough to show the far-reaching effect of "forward-induction" considerations to the deterrence strategy. To the best of my knowledge, this result is the first serious application of forward induction to a complete-information-game model that is widely known in applied economics, although there are several such results regarding incomplete-information games.

In the following of this introduction, I first review literature on "forward induction," after which I construct the model in section 2. In section 3, I define the solution concept extensive-form rationalizability. In this article, I adopt the formalization of extensive-form rationalizability as in Battigalli (1997). In section 4, I analyzed the game and state the result. Section 5 discusses the relationship between the current study and the closely related research. Section 6 concludes.

1.1.1 Related Literature

Forward induction is argued as a key factor of "strategically stable equilibrium" by Kohlberg and Mertens (1986). The "forward-induction outcomes" are informally described for specific examples,[1] and several solution concepts that captures forward induction reasoning are proposed by Kohlberg and Mertens (1986) and Reny (1992), among others. An attempt to directly define "forward induction outcome" is done by Govindan and Wilson (2009).

Applications of forward induction considerations to complete-information settings include Ben-Porath and Dekel (1992) and Osborne (1990), among others. Ben-Porath and Dekel (1992) considered an implication of costly signaling ("money burning") of intended strategies and show that, if (only) one player is allowed to costly signal their intention, the player's most preferred outcome is the unique outcome that survives the maximal iterative deletion of weakly dominated strategies. In addition to that the solution procedure used is different, our game with the participation stage differs from the "money-burning" game because the game with participation stage does not allow participation to the subsequent game without incurring a cost. The work by Osborne (1990) is closely related to the present research and will discussed in Section 5.

While the studies above have tried to characterize outcomes that are consistent with "forward induction" considerations, another strand of literature tries to directly formalize what constitutes "forward-induction reasoning." According forward-induction reasoning, players try to rationalize the other players' behavior as much as possible. The first of them is thought of as Pearce (1984)'s extensive-form rationalizability,[2] which I adopt here, too. Since Pearce's formulation was complicated and difficult to interpret, several studies tries to deepen our understanding toward the solution concept (Battigalli, 1997; Reny, 1992). As to incomplete-information games, Battigalli and Siniscalchi (2002, 2003)'s Δ-rationalizability provides a useful analytical tool to embody forward-induction reasoning and explicit description of the set of possible beliefs.

[1] For another example of forward induction in a complete-information game, see van Damme (1989).

[2] See Perea (2010) for a discussion on forward-induction reasoning and backward-induction reasoning.

1.2 Model

We consider the following two-player finite extensive-form game with three stages. The first stage is the participation stage: Players 1 and 2 simultaneously choose whether to participate (Y) in the subsequent game or not (N). Only when both of the players choose Y does the game proceed to the next stage. Participation to the subsequent game incurs participation cost of $\varepsilon > 0$, or equivalently, staying out from the game gives a player a payoff of ε.

The interpretation is either on investment on offensive-weapons (MAD) or construction of factories (self-enforcing collusion). Assume $0 < \varepsilon < 1$.

Table 1.1. Chapter 1: First Stage

	Out	In
Out	ε, ε	$\varepsilon, 0$
In	$0, \varepsilon$	$0, 0$ & *continue to the next stage.*

In the second stage, players play a Prisoners' Dilemma (PD) game. Players 1 and 2 simultaneously choose Cooperate (C) or Defect (D), the interpretation being initiation of direct attack (D) or avoidance of it (C) in MAD and collusive pricing (C) or secret price cutting (D) in collusion.

Table 1.2. Chapter 1: Second Stage

	C	D
C	$1, 1$	$-1, 2$
D	$2, -1$	$0, 0$

After both players choose actions, they observe the outcome in PD game. Then, in the third stage, the players play the following "War and Peace" game, in which players simultaneously choose Peace (P) or War (W).

Table 1.3. Chapter 1: Third Stage

	P	W
P	$0, 0$	$-\gamma, -\alpha$
W	$-\alpha, -\gamma$	$-\beta, -\beta$

We assume that $0 < \alpha < \beta < \gamma$. Hence the War and Peace game is a symmetric coordination game with Pareto-ranked stage game Nash equilibria. We could interpret this stage as the choice between initiating a (destructive) War or keep Peace, literally, or as triggering a price War (W) or avoidance of it (P). Another interpretation in terms of collusion example is a report to antitrust authority and ask for reduction of fines (so-called "Leniency Policy").

We assume that there is no discounting.

1.3 Extensive-Form Rationalizability

I apply the extensive-form rationalizability to derive a forward-induction result (Pearce, 1984). I follow the formulation as in Battigalli (1997).

We define extensive-form rationalizability for our finite extensive-form game. To do this, we need several notations.

A T-period two-player finite extensive-form game with complete information and observable actions is formulated as follows:[3]

$$(A_1, A_2, \mathcal{H}, \mathcal{Z}, u_1, u_2).$$

For $i = 1, 2$, A_i is the set of player i's actions. Following the usual convention, we denote by $-i$ the opponent of player i. $\mathcal{H}$ is the set of non-terminal histories $\emptyset$ (the initial history) or those of the form $(a^1, a^2, \ldots, a^t)$ for $1 \leq t \leq T - 1$, where $a^\tau = (a_1^\tau, a_2^\tau) \in A_1 \times A_2$ $(1 \leq \tau \leq t)$, is an action profile taken at period τ. I consider only a game with observable actions, so there are only public histories. $\mathcal{Z}$ is the set of terminal histories of the form $(a^1, a^2, \ldots, a^t)$ for $1 \leq t \leq T$.[4] For $i = 1, 2$, $u_i : \mathcal{Z} \to \mathbb{R}$ is player i's (extensive-form) payoff function.

For each non-terminal history $h \in \mathcal{H}$ and for $i = 1, 2$, let $A_i(h) \subset A_i$ be the set of player i's actions available at h. A player i's (pure) strategy is a function σ_i which maps $\mathcal{H}$ into A_i and satisfies $\sigma_i(h) \in A_i(h)$ for all $h \in \mathcal{H}$. Let Σ_i be the set of player i's strategies. A strategy profile $\sigma = (\sigma_1, \sigma_2)$ induces a unique terminal history. Denote by the *outcome function* $\zeta(\sigma) \in \mathcal{Z}$ the terminal history induced by strategy profile σ. A non-terminal history $h \in \mathcal{H}$ is *reached by strategy* σ (or σ *reaches* h) if the path $\zeta(\sigma)$ includes h. Similarly, $\sigma_i \in \Sigma_i$ *reaches* $h \in \mathcal{H}$ if

[3] In our game, $T = 3$.

[4] I allow the possibility that the game ends before period T.

there exists some $\sigma_{-i} \in \Sigma_{-i}$ such that (σ_i, σ_{-i}) reaches h. Let $\Sigma_i(h)$ be the set of strategies of player i which reaches h.

Using the outcome function ζ, we can define an equivalence relation $\sim_i$ on Σ_i by

$$\sigma_i \sim_i \sigma_i' \text{ if and only if } \zeta(\sigma_i, \sigma_{-i}) = \zeta(\sigma_i', \sigma_{-i}) \text{ for all } \sigma_{-i} \in \Sigma_{-i}$$

Note that since the terminal histories are identified with the path of action profiles, $\sigma_i \sim_i \sigma_i'$ if and only if σ_i and σ_i' differ only at histories which neither σ_i nor σ_i' reaches. The quotient space $\Sigma_i / \sim_i$ is called the set of *plans of actions*. In the definition of extensive-form rationalizability, the difference between strategies and plans of actions does not affect the outcome. While we use plans of actions in the subsequent analysis, here we use the word *strategies* to simplify the terminology.

The test of rationalizability consists in checking *optimality* of some strategy against some *belief* over the opponent's strategies. Now we define these notions in turn.

Definition 1 (Conditional Probability System (Rênyi, 1955)) A *conditional probability system (CPS)* of player i about the opponent's strategies is a function μ that maps $h_i \in \mathcal{H}_i$ to $\mu(\cdot|h_i) \in \Delta(\Sigma_{-i})$,[5] satisfying

Regularity: $\mu(\cdot|h_i)$ is a probability measure over σ_{-i} for any h_i,

Properness: $\mu(\Sigma_{-i}(h_i)|h_i) = 1$ for any h_i, and

Consistency: μ satisfies Bayes' law on histories to which μ attaches positive probability in preceding nodes.

Let $\mathcal{C}(\Sigma_{-i})$ be the set of CPSs of player i.

To define *optimality*, define the strategic-form payoff function $U_i : \Sigma \to \mathbb{R}$ by

$$U_i(\sigma) := u_i(\zeta(\sigma)).$$

[5]For a given set Ω, $\Delta(\Omega)$ denotes the set of probability measures on Ω, where the σ-algebra is appropriately chosen.

A strategy $\sigma_i \in \Sigma_i$ is *optimal in $S_i \subset \Sigma_i$ against $m \in \Delta(\Sigma_{-i})$* if[6]

$$\sum_{\sigma_{-i}} U_i(\sigma_i, \sigma_{-i}) m(\sigma_{-i}) \geq \sum_{\sigma_{-i}} U_i(\sigma'_i, \sigma_{-i}) m(\sigma_{-i}) \text{ for all } \sigma'_i \in S_i.$$

A strategy σ_i is *sequentially rational against* $\mu_i \in \mathcal{C}(\Sigma_{-i})$ if at any history h_i such that σ_i reaches h_i, σ_i is optimal in $\Sigma_i(h_i)$ against $\mu_i(\cdot|h_i)$.

The procedure for extensive-form rationalizability requires that each strategy is "justified" by a CPS, whose support is included in the opponent's "justified" strategies.

Definition 2 (Extensive-Form Rationalizability) Let $\Sigma_i^0 := \Sigma_i$. For each $k \geq 0$, $\sigma_i \in \Sigma_i^{k+1}$ if and only if there is a $\mu_i \in \mathcal{C}(\Sigma_{-i})$ such that

1. σ_i is sequentially rational against μ_i.
2. If $\Sigma_{-i}(h_i) \cap \Sigma_{-i}^k \neq \emptyset$, then $\mu_i(\Sigma_{-i}^k|h_i) = 1$.

If $\sigma \in \Sigma^k$, then σ is said to be k-rationalizable. If $\sigma \in \Sigma^k$ for all k, then σ is said to be extensive-form rationalizable.

1.4 Analysis

1.4.1 Actions, Histories, and Strategies (Plans)

Recall that the sets of stage-game actions are $\{\text{Out}, \text{In}\}$, $\{C, D\}$, and $\{P, W\}$

The set of non-terminal histories is

$$\{\emptyset\} \cup \{\text{Out}, \text{In}\}^2 \cup \left(\{\text{In}\}^2 \times \{C, D\}^2\right),$$

and the set of terminal histories is

$$\{\text{Out}, \text{In}\} \cup \{\text{In}, \text{Out}\} \cup \{\text{Out}, \text{Out}\} \cup \left(\{\text{In}\}^2 \times \{C, D\}^2 \times \{P, W\}^2\right).$$

For example, a terminal history $(\text{In}, \text{In}, C, D, P, W)$ means a history (a path of play) in which both players choose In in the first stage, player 1 chooses C and player 2

[6] Here we use the notation when Σ_{-i} is finite, which is the case in our setting.

chooses D in the second stage, and player 1 chooses P and player 2 chooses W in the third stage. The payoff function u_1 for player 1 satisfies

$$\begin{aligned}
&u_1(\text{Out}, \text{In}) = u_1(\text{Out}, \text{Out}) = \varepsilon \\
&u_1(\text{In}, \text{Out}) = 0 \\
&u_1(a_1^2, a_2^2, a_1^3, a_2^3) = u_1^2(a_1^2, a_2^2) + u_1^3(a_1^3, a_2^3),
\end{aligned}$$

where u_1^2 and u_1^3 are player 1's payoff functions in Prisoners' Dilemma game and War and Peace game, respectively. The payoff function u_2 for player 2 is given in a similar manner.

A strategy is a function satisfying

$$\begin{aligned}
&\sigma(\emptyset) \in \{\text{Out}, \text{In}\} \\
&\sigma(\text{In}, \text{In}) \in \{C, D\}, \text{ and} \\
&\sigma(a_1^2, a_2^2) \in \{P, W\} \text{ for } (a_1^2, a_2^2) \in \{C, D\}^2.
\end{aligned}$$

The set of plans of actions is

$$\{\text{Out}\} \cup \{(\text{In}, a, b(C), b(D)) : \; a \in \{C, D\}, (b(C), b(D)) \in \{P, W\}^2\}.$$

where $b(C)$ and $b(D)$ prescribe the actions taken in response to the opponent's action C and D in Prisoners' Dilemma game, respectively. For notational convenience, I write $(\text{In}, a, b(C), b(D))$ as $\text{In}ab(C)b(D)$.

1.4.2 Subgame-Perfect Equilibrium

Observe that both (P, P) and (W, W) are Nash equilibria in the third stage. In particular, using the Pareto-dominated equilibrium (W, W) as a punishment, cooperation in Prisoners' Dilemma game is possible:

Proposition 1 Let $\beta > 1$. Then a strategy profile (σ_d, σ_d) of deterrence strategy σ_d such that

$$\begin{aligned}
&\sigma_d(\emptyset) = \text{In} \\
&\sigma_d(\text{In}, \text{In}) = C
\end{aligned}$$

$$\sigma_d(a_1^2, a_2^2) = \begin{cases} P & \text{if} \quad a_1^2 = a_2^2 = C \\ W & \text{otherwise} \end{cases}$$

is a subgame-perfect equilibrium (SPE).[7]

The plan of action InCPW corresponds to the deterrence strategy; thus we call the profile (InCPW, InCPW) a "deterrence SPE."

1.4.3 When The Deterrence SPE is Not Extensive-Form Rationalizable

Credibility of mutual punishment is cast doubt by the literature on renegotiation-proofness, which is first discussed by Bernheim and Ray (1989) and Farrell and Maskin (1989), among others. They argue that if players could communicate before they choose action in each period, they would avoid such Pareto-dominated outcome by *renegotiation*. Since defining *renegotiation-proof equilibria* in infinitely repeated games has conceptual difficulties there are several versions of renegotiation-proof equilibria.

In contrast, Benoît and Krishna (1993) formulated renegotiation-proofness in finite games.[8] In finite dynamic games which has a well-defined terminal history, there is little controversy on the formulation of renegotiation and definition of renegotiation-proof equilibrium.

Here I argue that when there is a participation stage, the deterrence SPE might not be extensive-form rationalizable. This result gives another reason why we could question the credibility of the mutual punishment and provides, to the best of my knowledge, the first economic example in which extensive-form rationalizability gives a different prediction from SPE in a complete-information game.

Theorem 1 Let $\beta > 1$. Suppose that the conditions

$$0 < \varepsilon < 1 \tag{1.1}$$

$$(2 - \varepsilon)(\beta - \alpha) > (2 - \alpha - \varepsilon)\gamma \tag{1.2}$$

[7] For the definition and expositions of SPE, please refer to textbooks such as Fudenberg and Tirole (1991), Myerson (1997), or Osbourne and Rubinstein (1994).

[8] Benoît and Krishna's renegotiation-proof equilibria are defined for finitely repeated games. It is straightforward to define the similar equilibrium concept for the current game.

are satisfied. Then a plan of action InCPW is not extensive-form rationalizable.

Moreover, if the additional condition

$$(1-\varepsilon)\beta \geq 2 \tag{1.3}$$

is satisfied, then the set of extensive-form rationalizable plans of actions is

$$\{\text{Out}, \text{In}CPP, \text{In}DPP, \text{In}DPW\}.$$

If the condition (1.3) is violated, then the set of extensive-form rationalizable plans of actions is {Out}.

Intuitively, the condition $\varepsilon > 0$ in (1.1) states that the participation cost (or the value of the outside option) is larger than the payoff which is obtained by DD and PP. Hence, combined with the assumption $\varepsilon < 1$, the value of the participation cost is intermediate in the sense that the decision of entering the game is justified only when a player believes that at least the payoff from CC is realized in the continuation game after entry. The second condition (1.2) states that the difference $\beta - \alpha$ should not be too small compared to the value γ. Moreover, if $2 - \alpha - \varepsilon > 0$,[9] it can be written as

$$\frac{2-\varepsilon}{2-\alpha-\varepsilon} > \frac{\gamma}{\beta-\alpha}.$$

The left-hand side is the ratio of the payoff from choosing P to choosing W when the player has unilaterally defected while the right-hand side is the ratio of the stage-game payoff from choosing P to choosing W conditional on the opponent choosing W. The condition states that compared to the relative loss from choosing P when the opponent is choosing W is smaller than the relative benefit from choosing P when the opponent is choosing P.

Example 1 For example, $\varepsilon = 0.5$, $\alpha = 0.5$, $\beta = 3.5$, and $\gamma = 5$ satisfies the conditions (1.1) and (1.2) but violates (1.3). If instead $\beta = 4.5$, then all the conditions in Theorem 1 are satisfied.

Proof of Theorem 1 To prove the result, I use the procedure proposed by Shimoji and Watson (1998) (hereafter SW). They prove that the procedure of maximal

[9] This is the case when my result makes a difference from Osborne (1990): see the Discussion section.

elimination of conditionally dominated strategy gives extensive-form rationalizable strategies.[10]

We briefly give the definitions and results in SW which are used in this paper. In our dynamic game, the players can condition their decisions on the observed histories of play, i.e., information sets. The set of (extensive-form) information sets is

$$\Lambda := \{\Sigma_1(h) \times \Sigma_2(h) : \ h = \emptyset, (\text{InIn}), \text{ or } (\text{In}, \text{In}, a_1, a_2) \text{ with } (a_1, a_2) \in \{C, D\}^2\},$$

which is common among the players. For example, an extensive-form information set at history (In,In,C,D) is

$$\{\text{In}CPP, \text{In}CPW, \text{In}CWP, \text{In}CWW\} \times \{\text{In}DPP, \text{In}DPW, \text{In}DWP, \text{In}DWW\}.$$

Let $X = X_i \times X_{-i}$ be a typical element of Λ.[11] A plan σ_i for player i is *strictly dominated in* X if $\sigma_i \in X_i$, $X_{-i} \neq \emptyset$, and there is a mixed plan $\hat{\sigma}_i \in \Delta X_i$ such that

$$U_i(\hat{\sigma}_i, \sigma_{-i}) > U_i(\sigma_i, \sigma_{-i}) \text{ for all } \sigma_{-i} \in X_{-i}.$$

The *conditional dominance* concept is a strict dominance concept on some information set.

Definition 3 (Conditional Dominance (Definition 3 in SW)) Given Λ and a set of profiles of plans $S \subset \Sigma$, a plan $\sigma_i \in \Sigma_i$ is conditionally dominated on (S, Λ) if there exists a set $X \in \Lambda$ such that σ_i is strictly dominated in $X \cap S$.

It is known that if we conduct the maximal iterative deletion of conditionally dominated plans, then the resulting set coincides with the set of extensive-form

[10] While my analysis deals with plans of actions, they are equivalent since each pair of plans of actions induces the same outcome as the corresponding pair of strategies.

[11] Note that in our setting, any element $X \in \Lambda$ can be written in the form $X = X_i \times X_{-i}$. This is an implication of perfect recall.

rationalizable plans (Theorem 3 in SW, p. 177).[12] [13]

Also, the following fact is used in the proof: there is no CPS whose support is included by S that justifies a plan if and only if the plan is conditionally dominated on (S, Λ) (Lemma 2 in SW, p. 177).

Since the game is symmetric, I prove the result for player 1's plans of actions.

The First Round of Conditional Dominance Procedure Consider the payoff matrix given in Table 1.4. The figure in each cell denotes the row player's payoff.

For each plan, we check whether the plan is deleted or not.

InCWP Consider a mixed strategy σ which choose Out with probability p and InCPP with probability $1-p$. Then the condition for which σ gives a higher payoff than InCWP against InCPP is

$$\varepsilon p + (1-p) > 1 - \alpha \tag{1.4}$$

$$\iff \frac{\alpha}{1-\varepsilon} > p. \tag{1.5}$$

Let

$$p := \frac{\alpha}{1-\varepsilon} - \eta \tag{1.6}$$

for $\eta > 0$. Note that the condition (1.5) is satisfied for $\eta > 0$.

[12] More precisely, denote

$$UD_i(S,\Lambda) := \{\sigma_i \in \Sigma_i : \ \sigma_i \text{ is not conditionally dominated on } (S,\Lambda)\}$$
$$UD(S,\Lambda) := UD_1(S,\Lambda) \times UD_2(S,\Lambda)$$

and recursively define

$$UD^0(S,\Lambda) := S$$
$$UD^{k+1}(S,\Lambda) := UD(UD^k(S,\Lambda),\Lambda)$$

for nonnegative integer k. Then Theorem 3 in SW states that $Q(\Lambda) := \cap_{k\geq 0} UD^k(\Sigma,\Lambda)$ coincides with the set of extensive-form rationalizable plans.

[13] Chen and Micali (2013) proved that the deletion procedure of conditional dominance is order-independent up to outcome equivalence.

Table 1.4. Chapter 1: First Round of Deletion

	Out	InCPP	InCPW	InCWP	InCWW	InDPP	InDPW	InDWP	InDWW
Out	ε	ε	ε	ε	ε	ε	ε	ε	ε
InCPP	0	1	1	$1-\gamma$	$1-\gamma$	-1	-1	$-1-\gamma$	$-1-\gamma$
InCPW	0	1	1	$1-\gamma$	$1-\gamma$	$-1-\alpha$	$-1-\alpha$	$-1-\beta$	$-1-\beta$
InCWP	0	$1-\alpha$	$1-\alpha$	$1-\beta$	$1-\beta$	-1	-1	$-1-\gamma$	$-1-\gamma$
InCWW	0	$1-\alpha$	$1-\alpha$	$1-\beta$	$1-\beta$	$-1-\alpha$	$-1-\alpha$	$-1-\beta$	$-1-\beta$
InDPP	0	2	$2-\gamma$	2	$2-\gamma$	0	$-\gamma$	0	$-\gamma$
InDPW	0	2	$2-\gamma$	2	$2-\gamma$	$-\alpha$	$-\beta$	$-\alpha$	$-\beta$
InDWP	0	$2-\alpha$	$2-\beta$	$2-\alpha$	$2-\beta$	0	$-\gamma$	0	$0-\gamma$
InDWW	0	$2-\alpha$	$2-\beta$	$2-\alpha$	$2-\beta$	$-\alpha$	$-\beta$	$-\alpha$	$-\beta$

The condition for which σ gives a higher payoff than InCWP against InCWP is

$$\begin{aligned} &\varepsilon p + (1-\gamma)(1-p) > 1-\beta \\ \iff &(\varepsilon - 1 + \gamma)p > \gamma - \beta. \end{aligned}$$

Substitute (1.6) into the above expression to get

$$(\varepsilon - 1 + \gamma)\left(\frac{\alpha}{1-\varepsilon} - \eta\right) > \gamma - \beta \tag{1.7}$$

$$\iff -\alpha + \gamma\frac{\alpha}{1-\varepsilon} - \eta(\varepsilon - 1 + \gamma) > \gamma - \beta. \tag{1.8}$$

Note that if

$$-\alpha + \gamma\frac{\alpha}{1-\varepsilon} > \gamma - \beta$$

is satisfied, then we can take $\eta > 0$ small enough that the expression (1.8) is satisfied. Then

$$\begin{aligned} &-\alpha + \gamma\frac{\alpha}{1-\varepsilon} > \gamma - \beta \\ \iff &(1-\varepsilon)(\beta - \alpha) > (1 - \alpha - \varepsilon)\gamma. \end{aligned}$$

Note that this condition is implied by the assumption (1.2). To show this, we prove the contrapositive: assume $(1-\varepsilon)(\beta-\alpha) \leq (1-\alpha-\varepsilon)\gamma$. Then

$$\begin{aligned} &(2-\varepsilon)(\beta-\alpha) \\ =&(1-\varepsilon)(\beta-\alpha) + \beta - \alpha \\ \leq&(1-\alpha-\varepsilon)\gamma + \beta - \alpha \\ =&(2-\alpha-\varepsilon)\gamma - (\gamma - \beta + \alpha) \\ <&(2-\alpha-\varepsilon)\gamma, \end{aligned}$$

as desired.

Now let σ be such that the above conditions are satisfied and $p > 0$ (i.e., take $\eta > 0$ small enough that the above conditions are met and $p > 0$).

Then, by construction, σ gives higher payoff to InCWP against InCPP, InCPW, InCWP, and InCWW. It is easily seen that this σ gives higher payoff to InCWP against InDPP, InDPW, InDWP, and InDWW since $\varepsilon > 0$ and $p > 0$.

InCWW Letting σ be a mixed strategy which choose N with probability p and InCPW with probability $1-p$, the proof follows analogously.

InDWP Let σ be a mixed strategy that chooses Out with probability p and InDPP with probability $1-p$. The condition for σ give a higher payoff than InDWP against InCPP is

$$\begin{aligned} & \varepsilon p + 2(1-p) > 2 - \alpha \\ \iff & \frac{\alpha}{2-\varepsilon} > p. \end{aligned}$$

Let

$$p := \frac{\alpha}{2-\varepsilon} - \eta \tag{1.9}$$

for $\eta > 0$. The condition for σ give a higher payoff than InDWP against InCPW is

$$\begin{aligned} & \varepsilon p + (2-\gamma)(1-p) > 2 - \beta \\ \iff & (\varepsilon - 2 + \gamma)p > \gamma - \beta \end{aligned}$$

Now substitute (1.9) into the above expression to yield

$$(\varepsilon - 2 + \gamma)\left(\frac{\alpha}{2-\varepsilon} - \eta\right) > \gamma - \beta \tag{1.10}$$

$$\iff -\alpha + \gamma\frac{\alpha}{2-\varepsilon} - \eta(\varepsilon - 2 + \gamma) > \gamma - \beta \tag{1.11}$$

Note that if

$$-\alpha + \gamma\frac{\alpha}{2-\varepsilon} > \gamma - \beta$$

is satisfied, then we can take $\eta > 0$ small enough that the expression

(1.11) is satisfied. Now

$$\begin{aligned} & -\alpha+\gamma\frac{\alpha}{2-\varepsilon} > \gamma-\beta \\ \iff & (2-\varepsilon)(\beta-\alpha) > (2-\varepsilon-\alpha)\gamma, \end{aligned}$$

which is exactly the assumption (1.2).

Let σ be such that the above conditions are satisfied and $p > 0$. Then, by construction, σ gives higher payoff to InDWP against InCPP, InCPW, InCWP, and InCWW. It is easily seen that this σ gives higher payoff to InDWP against InDPP, InDPW, InDWP, and InDWW since $\varepsilon > 0$ and $p > 0$.

InDWW Letting σ be a mixed strategy which choose Out with probability p and InDPW with probability $1-p$, the proof follows analogously.

The Other Plans The other plans are not deleted. Indeed, Out is justified by a CPS μ such that[14]

$$\mu(\text{Out}|\emptyset) = 1.$$

InCPP is justified by a CPS μ such that

$$\mu(\text{Out}|\emptyset) = 0$$

$$\mu(\text{In}ab(C)b(D)|\emptyset) = \begin{cases} 1 \text{ if } (a, b(C), b(D)) = (C, P, W) \\ 0 \text{ otherwise} \end{cases}$$

$$\mu(\text{In}ab(C)b(D)|\text{In}, \text{In}) = \begin{cases} 1 \text{ if } (a, b(C), b(D)) = (C, P, W) \\ 0 \text{ otherwise} \end{cases}$$

$$\mu(\text{In}ab(C)b(D)|\text{In}, \text{In}, C, C) = \begin{cases} 1 & \text{if } (a, b(C), b(D)) = (C, P, W) \\ 0 & \text{otherwise.} \end{cases}$$

$$\mu(\text{In}ab(C)b(D)|\text{In}, \text{In}, C, D) = \begin{cases} 1 & \text{if } (a, b(C), b(D)) = (D, P, P) \\ 0 & \text{otherwise.} \end{cases}$$

[14] Note that the sequential rationality condition uses the values of μ only at histories which are reached by the plan.

InCPW is justified by a CPS μ which is the same in the case of InCPP except that

$$\mu(\text{In}ab(C)b(D)|\text{In},\text{In},C,D) = \begin{cases} 1 & \text{if } (a,b(C),b(D)) = (D,W,W) \\ 0 & \text{otherwise.} \end{cases}$$

InDPP is justified by a CPS μ such that

$$\mu(Out|\emptyset) = 0$$

$$\mu(\text{In}ab(C)b(D)|\emptyset) = \begin{cases} 1 \text{ if } (a,b(C),b(D)) = (C,P,P) \\ 0 \text{ otherwise} \end{cases}$$

$$\mu(\text{In}ab(C)b(D)|\text{In},\text{In}) = \begin{cases} 1 \text{ if } (a,b(C),b(D)) = (C,P,P) \\ 0 \text{ otherwise} \end{cases}$$

$$\mu(\text{In}ab(C)b(D)|\text{In},\text{In},D,C) = \begin{cases} 1 & \text{if } (a,b(C),b(D)) = (C,P,P) \\ 0 & \text{otherwise.} \end{cases}$$

$$\mu(\text{In}ab(C)b(D)|\text{In},\text{In},D,D) = \begin{cases} 1 & \text{if } (a,b(C),b(D)) = (D,P,P) \\ 0 & \text{otherwise.} \end{cases}$$

InDPW is justified by a CPS μ which is the same as in the case of InDPP except that

$$\mu(\text{In}ab(C)b(D)|\text{In},\text{In},D,D) = \begin{cases} 1 & \text{if } (a,b(C),b(D)) = (D,P,W) \\ 0 & \text{otherwise.} \end{cases}$$

The second round Now the remaining plans are

Table 1.5. Chapter 1: Second Round of Deletion

	Out	InCPP	InCPW	InDPP	InDPW
Out	ε	ε	ε	ε	ε
InCPP	0	1	1	-1	-1
InCPW	0	1	1	$-1-\alpha$	$-1-\alpha$
InDPP	0	2	$2-\gamma$	0	$-\gamma$
InDPW	0	2	$2-\gamma$	$-\alpha$	$-\beta$

It is easily seen that InCPW is conditionally dominated by InCPP (take $\Sigma_1((\text{In}, \text{In}, C, D)) \times \Sigma_2((\text{In}, \text{In}, C, D))$ as X). The other plans Out, InCPP, InDPP, and InDPW are justified by the same CPSs as in the first round.

The Third Round Now the remaining plans are

Table 1.6. Chapter 1: Third Round of Deletion

	Out	InCPP	InDPP	InDPW
Out	ε	ε	ε	ε
InCPP	0	1	-1	-1
InDPP	0	2	0	$-\gamma$
InDPW	0	2	$-\alpha$	$-\beta$

The plans Out, InDPP, and InDPW are again justified by the same CPSs as in the first round.

To derive the condition under which InCPP is justified, consider a mixed strategy of the opponent which chooses Out with probability 0, InCPP with probability p, InDPP with probability q, and InDPW with probability $1 - p - q$.[15] Then the sequential rationality condition at $\emptyset$ requires that following InCPP gives a higher payoff than deviating to Out:

$$2p - 1 \geq \varepsilon$$
$$\iff p \geq \frac{\varepsilon + 1}{2}.$$

The sequential rationality condition at (In, In) requires that following InCPP gives a higher payoff than deviating to InDPP,InDPW,InDWP, or InDWW.[16]

[15]Note that the sequential rationality condition of InCPP at $\emptyset$ is least restrictive if the probability of the opponent choosing Out is 0, that is, whenever InCPP is sequentially rational against a CPS with $\mu(\text{Out}|\emptyset) > 0$, there is a CPS μ' with $\mu'(\text{Out}|\emptyset) = 0$ against which InCPP is sequentially rational. This is because the possibility of the opponent choosing Out with positive probability only reduces the expected payoff from entering. The rigorous proof is simple, but is available from the author upon request.

[16]Deviation to InCPW, InCWP, or InCWW does not increase the expected payoff because the un-improvability conditions for those deviations are implied by the sequential rationality conditions at histories $(\text{In}, \text{In}, C, C)$ and $(\text{In}, \text{In}, C, D)$. In particular, in the current setting, it is easily seen that after choosing C, choosing P is always optimal since with probability one, the opponent reacts to C by P.

The deviations to InDPP, InDPW, InDWP, and InDWW yields

$$2p + (1 - p - q)(-\gamma),$$

$$2p + (-\alpha)q + (1 - p - q)(-\beta),$$

$$(2 - \alpha)p + (1 - p - q)(-\gamma),$$

and

$$(2 - \alpha)p + (-\alpha)q + (1 - p - q)(-\beta),$$

respectively. Note that the deviation to InDWP is inferior to that to InDPP and the deviation to InDWW is inferior to that to InDPW. Hence the conditions which must be satisfied are $p \geq \frac{\varepsilon+1}{2}$ and

$$\begin{cases} 2p - 1 \geq 2p + (1 - p - q)(-\gamma) \\ 2p - 1 \geq 2p + (-\alpha)q + (1 - p - q)(-\beta) \end{cases}$$
$$\iff \begin{cases} \gamma \geq 1 + \gamma(p + q) \\ \beta \geq 1 + \beta p + (\beta - \alpha)q \end{cases}.$$

Note that the right-hand sides of the two inequalities are increasing in p and q. Hence we can let $p = \frac{\varepsilon+1}{2}$ and $q = 0$. Then the two inequalities become

$$\begin{cases} \beta \geq 1 + \frac{\beta(\varepsilon+1)}{2} \\ \gamma \geq 1 + \frac{\gamma(\varepsilon+1)}{2} \end{cases}$$
$$\iff \begin{cases} (1 - \varepsilon)\beta \geq 2 \\ (1 - \varepsilon)\gamma \geq 2 \end{cases}$$

The former condition implies the latter. Therefore, InCPP is justifiable if and only if $(1 - \varepsilon)\beta \geq 2$.

If the condition is satisfied, there is no plan which is deleted. Otherwise, InCPP is deleted and the deletion proceeds to the next round.

The Forth Round The remaining plans are

	Out	InDPP	InDPW
Out	ε	ε	ε
InDPP	0	0	$-\gamma$
InDPW	0	$-\alpha$	$-\beta$

Since $\varepsilon > 0$, both InDPP and InDPW are strictly dominated by Out.

□

1.5 Discussion: Relationship with Osborne (1990)

Osborne (1990) gives a forward induction result in finitely repeated games. He shows that if a player's deviation from a candidate equilibrium path is interpreted as an unambiguous message to deviate to another path of play and if the other player also has an incentive to follow the suggested path, then the candidate path is not a part of stable equilibrium in the sense of Kohlberg and Mertens (1986). To compare my work with his, I give the version of his proof conformed to my model. Denote by $B_i := (B_{i,i}, B_{i,-i})$ and $C_i := (C_{i,i}, C_{i,-i})$ the best and the second-best outcomes for player i in the third stage. Recall that the payoff functions at the second and third stages are given by u_i^2 and u_i^3, respectively.

Proposition 2 (Osborne (1990)) Consider a path of play $(\text{In}, \text{In}, a_1, a_2, b_1, b_2)$ in the game. If there is a deviation at the second stage[17] $\tilde{a}_i$ for player i such that

$$u_i^2(\tilde{a}_i, a_{-i}) + u^3(C_i) < u_i^2(a_i, a_{-i}) + u_i^3(b_i, b_{-i}) < u_i^2(\tilde{a}_i, a_{-i}) + u_i^3(B_i) \tag{1.12}$$

and

$$B_{i,-i} \text{ is player } -i\text{'s unique best response to } B_{i,i}, \tag{1.13}$$

then the path of play $(\text{In}, \text{In}, a_1, a_2, b_1, b_2)$ is not stable.

[17]Since the payoff from a deviation from Y to N in the participation stage is independent of the other player's response, only a deviation at the second stage gives a sensible comparison.

The condition (1.12) states that player i can gain by deviating from a_i to $\tilde{a}_i$ only if the deviation results in B_i in the third stage. The condition (1.13) states that player $-i$ cannot gain by deviating from the outcome B_i in the third stage, which was "suggested" by player i. Note that the condition (1.13) is always satisfied in our game since the War and Peace game is a coordination game.

Now, under some parameter values, Osborne (1990)'s result has a bite to the stability of the path $(\text{In}, \text{In}, C, C, P, P)$.

Proposition 3 If $\alpha > 1$, then the path $(\text{In}, \text{In}, C, C, P, P)$ is not stable.

Proof of Proposition 3 By symmetry, consider player 1. Then the deviation from $a_1 = C$ in the second stage is $\tilde{a}_1 = D$ and the outcomes B_1 and C_1 are

$$B_1 = (P, P) \text{ and}$$
$$C_1 = (W, P),$$

respectively. Now the values of the three terms are

$$u_i^2(\tilde{a}_i, a_{-i}) + u^3(C_i) = 2 - \alpha$$
$$u_i^2(a_i, a_{-i}) + u_i^3(b_i, b_{-i}) = 1$$
$$u_i^2(\tilde{a}_i, a_{-i}) + u_i^3(B_i) = 2.$$

Now it is easily seen that the condition (1.12) is satisfied when $\alpha > 1$. □

Intuitively, when player 1 deviates from C to D, the deviation unambiguously signals that player 1 is not intending to achieve an outcome other than (P, P) in the third stage. Given that player 1 intends to play P, player 2 has no incentive to play W.

On the other hand, if $\alpha \leq 1$, then player 1 is better off by the deviation to D in the second stage as long as player 1 believes that (P, P) or (W, P) is achieved in the third stage. This signal does not convey a unique action that player 1 intends to play in the third stage; in particular, player 2 cannot exclude the possibility that player 1 might play W in the third stage. This prevents the result of Osborne from being applicable to the path $(\text{In}, \text{In}, C, C, P, P)$.[18]

[18]Note that Osborne (1990)'s result only gives a *sufficient* condition for a path to be unstable.

1.6 Conclusion

In this paper I showed that, under certain circumstances, a subgame-perfect equilibrium in which a cooperation is supported by a threat of mutual punishment might not be extensive-form rationalizable.

My result here *does not* state that the efficient outcome cannot be a plausible outcome under our assumptions. As Battigalli and Friedenberg (2012) states, a subgame-perfect equilibrium outcome always constitutes an *extensive-form best response set*, which corresponds to a set of Δ-rationalizable[19] outcomes *for some belief restriction* Δ. Thus our result states that if no restriction is imposed on the players' beliefs, the forward-induction reasoning excludes the celebrated deterrence equilibrium, which has the same structure as the grim trigger strategy profile in infinitely repeated Prisoners' Dilemma games.

While the literature in renegotiation-proof equilibria argue that players might avoid Pareto-dominated punishment if they could renegotiate, this study suggest that an implicit message of participating in the game also might convey the deviator's unwillingness to mutually punish. Even in a simple economic example, our model suggests a possibility that the "forward-induction reasoning" (embodied in extensive-form rationalizability) might have interesting predictions that are not obtained by subgame-perfect equilibrium.

A promising future research would be an extension of our current result to infinitely repeated prisoners' dilemma game under perfect monitoring.

[19] A rationalizability procedure in which players' possible CPS's are restricted by a set of belief restrictions $\Delta = (\Delta_i)_i$.

Chapter 2 Credibility of Deterrence and Extensive-Form Rationalizability: Infinite-Horizon Case

2.1 Introduction

The problem on credibility of deterrence by a threat of mutual punishment is studied in Isogai (2016) using a simple, finite-horizon model of strategic interactions. While the model captures the idea that forward-induction considerations might have an unexpected effect, actual strategic relationships often involve long-term strategic interactions as in industrial cartels or interactions between nations, which would last unless disrupted by some exit events. In this paper, I study a natural extension of the model in Isogai (2016) to infinite-horizon dynamic game in which the game might continue forever, but the strategic relationship might be terminated by the event of mutual punishment or voluntary exit. Examples include industrial cartels, in which the punishment is a leniency application, and the model of mutual assured destruction (MAD), in which the punishment is a warfare employing the massive destructive weapons.

The model is an infinitely-repeated game augmented with entry choice in the initial period and with exit options in each stage game.[1] Each stage game consists of two phases. In the first phase, players play the prisoners' dilemma game. In

[1]Therefore, strictly speaking, the game should be referred to as an infinite-horizon dynamic game.

the second phase, players choose whether to continue the game or choose among two kinds of exit options: one with and without mutually destructive punishment. Unless any exit or punishment occurs, the game continues indefinitely. Each stage game payoff is discounted by a common discount factor. We assume that the players are not patient enough to sustain cooperation with the threat of mutual defection in prisoners' dilemma game but is patient enough to sustain cooperation if players can credibly use punishment option as a threat. Then perfect cooperation (i.e., the outcome in which players cooperate in the prisoners' dilemma game and do not exit the game) is supported only by a strategy profile in which players punish each other if a deviation in the prisoners' dilemma game is observed. As the solution concept employed is subgame-perfect equilibrium, the strategy profile is regarded as self-enforcing in the traditional view. My result, however, shows that, under some conditions, this strategy profile is not strategically stable in the sense of extensive-form rationalizability (Pearce, 1984), which is regarded as one that embodies forward-induction considerations.

The intuition of the result is similar to the previous work. Suppose that a deviation from cooperation has occurred. Then the players would decide whether or not to trigger punishment. If players believe that each is rational and the deviation was deliberate, each realizes that choosing to punish after deviation is never optimal against any belief the deviator might hold. Indeed, for any belief, either staying out from the game or exiting without punishment is a better choice. Having concluded that the deviator would not rationally choose punishment, since the punishment is mutually detrimental, both players would not punish each other, leaving the deviator unpunished. This is in contradiction to the punishment being the deterrence of deviation.

This result shows that perfect cooperation is not strategically stable under certain conditions. Moreover, we can show that since the exit option without punishment also serves as an outside option, at any period, a unilateral deviation from cooperation is not followed by mutual punishment.

Not only as an extension from the previous model, the analysis of infinite-horizon model has important implications to economic modeling using game theory. In the equilibrium analysis of infinite-horizon dynamic games, vast number of outcomes are supported by subgame-perfect equilibria and it seems almost impossible to choose one of those as a plausible outcome of the game. This study shows that using the tool of epistemic game theory, we can restrict attention to some strategically stable outcomes if we are certain that some conditions are met.

Another important implication of our work is that in real cartels, leniency applications (or severe price war) might not be a credible threat to deter price cutting by cartel firms. Indeed, it is frequently observed that even when a price cutting and business stealing by a cartel firm is known to other cartel members, the deviant firm is not punished immediately.[2] This behavior is not consistent with the theory of optimal deterrence in repeated games and is regarded as a puzzle in studies of cartels. Bernheim and Madsen (2017) explain this phenomenon as equilibrium outcome when the cartel firms compete in the Bertrand fashion, they are not patient enough to sustain full collusion, and they have cost advantages in their home markets. My result here gives another explanation to the phenomenon from the viewpoint of epistemic game theory.

The paper is structured as follows: in section 2, I construct the model. In section 3, I define the collection of histories, payoff functions, and strategies. In section 4, I give the formal definition of solution concept which I use in this paper. Section 5 analyzes the model and give the result.

2.2 Model

We consider an infinite-horizon model which is a natural extension of the one considered in Isogai (2016). There are two players $i = 1, 2$ and the time goes indefinitely $t = 0, 1, 2, \ldots$. As is conventional, when I denote a player by i, the opponent of i is denoted by $-i$. In the first period $t = 0$, the players simultaneously decide whether or not to enter the strategic interaction. We normalize the value of outside option to 0 and assume that the entry cost in the initial period is $\varepsilon > 0$. From period $t = 1$ on, the dynamic game consisting of an extensive-form stage game is played. The stage game has two phases:

[2]On the historical record of observed deviations, see also Bernheim and Madsen (2017).

Phase A The players play the following prisoners' dilemma game:

Table 2.1. Chapter 2: Phase A

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

where $g, l > 0$.

Phase B After observing the outcome in the prisoners' dilemma game, players choose an action among Tolerate (T), Exit (E), or War (W) with the following payoffs:

Table 2.2. Chapter 2: Phase B

	T	E	W
T	$0, 0$	η, η	$-\gamma, -\alpha$
E	η, η	η, η	$-\gamma, -\alpha$
W	$-\alpha, -\gamma$	$-\alpha, -\gamma$	$-\beta, -\beta$

where $\varepsilon \geq \eta \geq 0$ and $\gamma > \beta > \alpha > 0$. η is interpreted as the value of the outside option. Note that by the assumptions on the values of α, β, and γ, the best response against W is W.

If E or W is played by at least one player, then the whole game is terminated in the period. If both players choose T, then the game proceeds to the next period. Hence continuation of the game reveals to both players that both have chosen T.

In the following, we call phase A in period t by stage t-A and likewise for phase B. Let us denote the phase-payoff functions by u_A and u_B, which is given by the payoff matrices above. The stage-game payoff is the sum $u_A + u_B$. Each stage-game payoff is discounted by the common discount factor $\delta \in [0, 1)$ and the payoffs for the whole dynamic game is given by the (expected) discounted sum of stage-game payoffs. The formal specification of the payoff function is given later, when I define the set of outcomes.

Let $A^0 := \{\text{In}, \text{Out}\}$ be the action set available to the players at period 0. At period $t = 1, 2, \dots$, the actions available to the players are given by

$$A := \{C, D\}$$

for stage A and

$$B := \{T, E, W\}$$

for stage B.

2.3 History, Payoff Function, and Strategy

The set of *non-terminal histories* (*decision nodes*) consists of the following sets:

$$\mathcal{H}^0 := \{\emptyset\}$$
$$\mathcal{H}_A^t := \{((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \mathbf{a}^2, \mathbf{b}^2, \dots, \mathbf{a}^{t-1}, \mathbf{b}^{t-1}) : \ \mathbf{b}^\tau = (T, T) \text{ for } 1 \leq \tau \leq t-1\}$$
$$\mathcal{H}_B^t := \{((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \mathbf{a}^2, \mathbf{b}^2, \dots, \mathbf{a}^{t-1}, \mathbf{b}^{t-1}, \mathbf{a}^t) : \ \mathbf{b}^\tau = (T, T) \text{ for } 1 \leq \tau \leq t-1\},$$

where $\emptyset$ denotes the initial history. Note that the actions at phase B must be (T, T) to continue the game. Denote

$$\mathcal{H}_A := \bigcup_{t=1,2,\dots} \mathcal{H}_A^t$$
$$\mathcal{H}_B := \bigcup_{t=1,2,\dots} \mathcal{H}_B^t$$
$$\mathcal{H} := \mathcal{H}^0 \cup \mathcal{H}_A \cup \mathcal{H}_B.$$

$\mathcal{H}_A$ and $\mathcal{H}_B$ are the sets of non-terminal histories at phases A and B, respectively. $\mathcal{H}$ is the set of all non-terminal histories.

The set of *terminal histories* (*outcomes*) consists of the following:

$$\mathcal{Z}^0 := \{(\text{Out}, \text{Out}), (\text{In}, \text{Out}), (\text{Out}, \text{In})\}$$
$$\mathcal{Z}^t := \{((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \mathbf{a}^2, \mathbf{b}^2, \dots, \mathbf{a}^t, \mathbf{b}^t) : \ \mathbf{b}^\tau = (T, T), \ 1 \leq \tau \leq t-1; \ \mathbf{b}^t \neq (T, T)\}$$
$$\mathcal{Z}^\infty := \{((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \mathbf{a}^2, \mathbf{b}^2, \dots) : \ \mathbf{b}^\tau = (T, T) \text{ for all } \tau\}$$

$\mathcal{Z}^t, \ t = 0, 1, 2, \dots,$ is the set of terminal histories in which the game ends at period t. $\mathcal{Z}^\infty$ is the set of terminal histories in which the game continues forever. Define

$$\mathcal{Z} := \left(\bigcup_{t=0,1,\dots} \mathcal{Z}^0 \right) \cup \mathcal{Z}^\infty$$

I use the terminologies "terminal histories" and "outcomes" interchangeably.

Also define the set of all histories by $\bar{\mathcal{H}} := \mathcal{H} \cup \mathcal{Z}$. Note that if a history is of the form $h = (\mathbf{c}^0, \mathbf{c}^1, \mathbf{c}^2 \ldots, \mathbf{c}^t)$, then any history h' of the form $h' = (\mathbf{c}^0, \mathbf{c}^1, \mathbf{c}^2 \ldots, \mathbf{c}^\tau)$ for $\tau < t$ is also a history. In this case, we say that h *passes through* h' or h' *is a predecessor of* h.

The payoff function $\tilde{U}_i : \mathcal{Z} \to \mathbb{R}$ for the dynamic game is given by

$$\tilde{U}_i(z) = \begin{cases} 0 & \text{if } z \in \{(\text{Out}, \text{In}), (\text{Out}, \text{Out})\} \\ -\varepsilon & \text{if } z = (\text{In}, \text{Out}) \\ -\varepsilon + (1-\delta) \sum_{\tau=1}^{t} \delta^{\tau-1} \{u^A(\mathbf{a}^\tau) + u^B(\mathbf{b}^\tau)\} & \text{if } z \in \mathcal{Z}^t \\ -\varepsilon + (1-\delta) \sum_{\tau=1}^{\infty} \delta^{\tau-1} \{u^A(\mathbf{a}^\tau) + u^B(\mathbf{b}^\tau)\} & \text{if } z \in \mathcal{Z}^\infty \end{cases}$$

A strategy is defined as a mapping σ from $\mathcal{H}$ into $A^0 \cup A \cup B$ such that

$$\sigma(h) \begin{cases} \in A^0 & \text{if } h = \emptyset \\ \in A & \text{if } h \in \mathcal{H}_A \\ \in B & \text{if } h \in \mathcal{H}_B \end{cases}$$

Let Σ_i be the set of strategies of player i and $\Sigma := \Sigma_i \times \Sigma_{-i}$.

Once a strategy profile $\sigma = (\sigma_i, \sigma_{-i})$ is given, the outcome $\zeta(\sigma) \in \mathcal{Z}$ is uniquely determined. Then the *strategic-form payoff function* $U : \Sigma \to \mathbb{R}$ is defined as the composition $U := \tilde{U} \circ \zeta$.

Since the players receive information about the opponent's strategy from the fact that the game has reached a certain history, we need the notation for the sets of possible strategies at histories. Let $h \in \mathcal{H}$ be some non-terminal history. Then define the set $\Sigma(h)$ of the strategy profiles which are *consistent with history* h by

$$\Sigma(h) := \{\sigma : \ \zeta(\sigma) \text{ passes through history } h\}.$$

Also let

$$\Sigma_i(h) := \{\sigma_i : \text{ there exists a } \sigma_{-i} \text{ such that } (\sigma_i, \sigma_{-i}) \in \Sigma(h))\}.$$

We say that σ_i *is consistent with* h if $\sigma_i \in \Sigma_i(h)$.

2.4 Extensive-Form Rationalizability

To formally define the concept of extensive-form rationalizability and Δ-rationalizability, we need to define the system of beliefs about the opponent's strategies.

Definition 4 (Conditional Probability System (Rényi, 1955)) Let $\mathcal{B}_i$ be the collection of Borel sets of Σ_i.[3] A conditional probability system (CPS) of player i is a mapping $\mu_i : \mathcal{B}_{-i} \times \mathcal{H} \to [0,1]$, where I denote $\mu_i(B,h)$ by $\mu_i(B|h)$, satisfying the following three properties:

Regularity For any h, $\mu_i(\cdot|h)$ is a probability measure on $\mathcal{B}_{-i}$.

Properness $\mu_i(\Sigma_{-i}(h)|h) = 1$ for any $h \in \mathcal{H}$.

Consistency For h and h', $\mu_i(B|h) = \mu_i(B|h')\mu_i(\Sigma_{-i}(h')|h)$.

Denote by $\mathcal{C}$ the set of CPSs.

Rationality stipulates that players maximize their expected payoffs given their beliefs.

Definition 5 (Sequential Rationality) Player i's strategy σ_i is called sequentially rational against $\mu_i \in \mathcal{C}$ if and only if at any history h for which $\sigma_i \in \Sigma_i(h)$,

$$\int U_i(\sigma_i, \sigma_{-i})\mu_i(d\sigma_{-i}|h) \geq \int U_i(\sigma_i', \sigma_{-i})\mu_i(d\sigma_{-i}|h)$$

for any $\sigma_i' \in \Sigma_i(h)$.

In Δ-rationalizability, we impose restrictions $\Delta = (\Delta_i)_i \subset \mathcal{C}^2$ of beliefs about the opponent's strategies.

Definition 6 (Δ-rationalizability) Let Δ be given. Consider the following procedure:

1. Let $\Sigma_i(\Delta, 0) := \Sigma_i$

[3]The topology on Σ_i is the metric topology constructed analogously as in Battigalli (2003). The metric thus constructed is complete and separable.

2. For $t \geq 0$ and $\Sigma(\Delta, t) = \Sigma_i(\Delta, t) \times \Sigma_{-i}(\Delta, t)$ given, $\sigma_i \in \Sigma_i(\Delta, t+1)$ if and only if there exists a $\mu_i \in \Delta_i$ such that

   (a) σ_i is sequential rational against μ_i

   (b) If at a history h, $\Sigma_{-i}(h) \cap \Sigma_{-i}(\Delta, t) \neq \emptyset$ holds, then μ_i must satisfy $\mu_i(\Sigma_{-i}(\Delta, t)|h) = 1$.

Then we say that $\Sigma(\Delta) := \cap_{t \geq 0} \Sigma(\Delta, t)$ is the set of Δ-rationalizable strategies.

If there is no restriction (i.e., $\Delta_i = \mathcal{C}$), then the Δ-rationalizability coincides with extensive-form rationalizability.

Our game falls in a "simple game" as in Battigalli (2003) and thus the existence of extensive-form rationalizable set is assured.

When we derive extensive-form (and Δ-) rationalizable sets, the concept of conditional dominance (Shimoji and Watson, 1998) is useful. Let $\Lambda = \{\Sigma_i(h) \times \Sigma_{-i}(h) : \ h \in \mathcal{H}\}$ be the collection of (extensive-form) information sets at non-terminal histories. We say that a strategy $\sigma_i \in \Sigma_i$ is *strictly dominated in* $X = X_i \times X_{-i} \in \Lambda$ if and only if $X_{-i} \neq \emptyset$ and there exists a mixed strategy $\sigma'_i \in \mathcal{P}(X_i)$[4] such that for any $\sigma_{-i} \in X_{-i}$,

$$U_i(\sigma'_i, \sigma_{-i}) > U_i(\sigma_i, \sigma_{-i}).$$

Unless otherwise mentioned, we use the word "dominance" for strict dominance. The conditional dominance concept requires that a strategy be strictly dominated at *some* information set:

Definition 7 (Conditional Dominance (Shimoji and Watson, 1998)) Given the collection Λ of information sets and a subset $S \subset \Sigma$, we say that $\sigma_i \in S$ is conditionally dominated in (S, Λ) if and only if there exists an $X \in \Lambda$ such that σ_i is strictly dominated in $X \cap S$.

The following easily follows from the definition.

Lemma 1 If σ_i is conditionally dominated in $(\Sigma(\Delta, t), \Lambda)$, then σ_i is never sequentially optimal against μ_i which satisfies $\Sigma_{-i}(\Delta, t) \cap \Sigma_{-i}(h) \neq \emptyset \implies \mu_i(\Sigma_{-i}(\Delta, t)|h) = 1$.

[4]For some (countable) set A, $\mathcal{P}(A)$ is the set of probability measures whose supports are in A.

Hence, at each Δ-rationalizability procedure, if a strategy is conditionally dominated, then it is deleted.

2.5 Results on Implausibility of Cooperation

I present a result that, under certain parameter values, the perfectly cooperative outcome (the outcome $((\mathrm{In}, \mathrm{In}), (C, C), (T, T), (C, C), (T, T), \ldots))$ does not pass the extensive-form rationalizability criterion.

To show that, first assume that the value of the discount factor δ is intermediate in the sense that it is not as high as to support cooperation with the threat of mutual defection in the prisoners' dilemma game but is high enough to support cooperation with the threat of the action W.

Specifically, I assume that δ satisfies

$$\delta_2 := \max\left\{0, \frac{g-\beta}{1+g-\beta}\right\} < \delta < \frac{g}{1+g} =: \delta_1 \tag{2.1}$$

We have the following standard result on subgame-perfect equilibrium:

Proposition 4 Under the assumption (2.1), the outcome path

$$((\mathrm{In}, \mathrm{In}), (C, C), (T, T), (C, C), (T, T), \ldots)$$

is supported in subgame-perfect equilibrium (SPE)[5] only by the symmetric strategy profile (σ_D, σ_D) such that, for $h_A \in \mathcal{H}_A$ and $h_B \in \mathcal{H}_B$,

$$\sigma_D(\emptyset) = \mathrm{In}$$

$$\sigma_D(h_A) = \begin{cases} C & \text{if } h_A = ((\mathrm{In}, \mathrm{In}), (C, C), (T, T), (C, C), \ldots, (T, T)) \\ D & \text{otherwise} \end{cases}$$

$$\sigma_D(h_B) = \begin{cases} T & \text{if } h_B = ((\mathrm{In}, \mathrm{In}), (C, C), (T, T), (C, C), \ldots, (C, C)) \\ W & \text{otherwise} \end{cases}$$

We call the strategy σ_D "deterrence strategy."

[5] For the definition and expositions of SPE, please refer to textbooks such as Fudenberg and Tirole (1991), Myerson (1997), or Osbourne and Rubinstein (1994).

While the perfectly cooperative outcome is supported in SPE by the deterrence strategy profile, the strategy profile might not be robust to forward-induction considerations. Intuitively, under some parameter values, a strategy which plays D in stage 1-A followed by the punishment (W) is strictly dominated by some (mixed) strategy. In other words, whatever belief a player who is intending to play D might hold, he could always do better by either staying out or exit without punishment. Therefore, if a player believes that her opponent is rational and never plays dominated strategies, the player would not punish the opponent.

Our first result is given as follows:

Proposition 5 Under the condition

$$(1-\delta)(1+g+\eta)-\varepsilon<0 \tag{2.2}$$

or

$$(1-\delta)(1+g-\beta)-\varepsilon<0 \quad \textbf{and} \quad \varepsilon>(1-\delta)\frac{(\gamma+\alpha+\eta-\beta)(1+g)-\alpha\gamma-\beta\eta}{\gamma+\alpha+\eta-\beta}, \tag{2.3}$$

then the path $((\text{In},\text{In}),(C,C),(T,T),\dots)$ is not supported by subgame-perfect equilibria which survive extensive-form rationalizability.

Example 2 It is not difficult to find the parameter values that satisfies the first condition (2.2). For the second condition (2.3), the parameter values $g=1$, $\delta=0.4$, $\varepsilon=0.5$, $\eta=0$, $\alpha=1$, $\beta=2$, and $\gamma=4$ are the ones which satisfy (2.3) while violating (2.2).

This result shows that if the infinite repetitions of the prisoners' dilemma game is endowed with the costly entry, the exit option, and the punishment option, the mutual punishment is not a credible way of deterring defection.

Proof of Proposition 5 We prove the proposition by showing that the deterrence strategy is not extensive-form rationalizable. To do this, we first show that the strategies that chooses In at $t=0$, plays D in the stage 1-A, and plays W after the history (D,C) are not extensive-form rationalizable. This is proved by showing that the strategy is dominated by some mixed strategy. Before showing this, observe that

1. If two strategies differ only at histories which the strategies are not consistent with, then they are outcome-equivalent. Formally, for two strategies σ_i and σ'_i, if $\sigma_i(h) \neq \sigma'_i(h)$ only for h such that $\sigma_i, \sigma'_i \notin \Sigma_i(h)$, then $\zeta(\sigma_i, \sigma_{-i}) = \zeta(\sigma'_i, \sigma_{-i})$ for all σ_{-i}.[6]

2. When player i chooses E or W at stage B, player $-i$'s choice between E and T at the same stage does not affect the payoff to the players. That is, if $(\sigma_i, \sigma_{-i}) \in \Sigma(h_B)$ and $\sigma_i(h_B)$ =E or W, then $U_i(\sigma_i, \sigma_{-i}) = U(\sigma_i, \sigma'_{-i})$ for σ_{-i} and σ'_{-i} such that

$$\begin{aligned}
&\sigma_{-i}(h_B) = E \\
&\sigma'_{-i}(h_B) = W \\
&\sigma_{-i}(h) = \sigma'_{-i}(h) \text{ for } h \neq h_B
\end{aligned}$$

I use the following notation to describe strategies: Out and $\text{In}a^1 b^1(C) b^1(D)$. The former denotes the class of strategies which chooses Out at $t = 0$. In the latter expression, $a^1 \in \{C, D\}$ denotes the player's choice at history $((\text{In}, \text{In}))$ and $b^1(C), b^1(D) \in \{T, E, W\}$ be the choice of the player at history $((\text{In}, \text{In}), (a^1, C))$ and $((\text{In}, \text{In}), (a^1, D))$, respectively.

Lemma 2 If we focus on player i's strategies σ_i such that

$$\begin{aligned}
&\sigma_i(h^0) = \text{Out, or} \\
&\sigma_i(h^0) = \text{In, } \sigma_i((\text{In}, \text{In})) = D \text{, and } \sigma_i(h_B^1) \in \{E, W\},
\end{aligned} \tag{2.4}$$

then it is enough to consider player i's strategies

$$\tilde{\Sigma}_i = \{\text{Out}, \text{InDEE}, \text{InDWE}, \text{InDEW}, \text{InDWW}\}$$

[6] For example, two strategies σ_i and σ'_i such that

$$\begin{aligned}
\sigma_i((\text{In}, \text{In})) &= \sigma'_i((\text{In}, \text{In})) = C \\
\sigma_i((\text{In}, \text{In}), (D, D)) &\neq \sigma'_i((\text{In}, \text{In}), (D, D)) \\
\sigma_i(h) &= \sigma'_i(h) \qquad \text{for } h \neq ((\text{In}, \text{In})), ((\text{In}, \text{In}), (D, D))
\end{aligned}$$

are outcome-equivalent.

and player $-i$'s strategies

$\tilde{\Sigma}_{-i} := \{\text{Out}, \text{InCEE}, \text{InCEW}, \text{InCWE}, \text{InCWW}, \text{InDEE}, \text{InDEW}, \text{InDWE}, \text{InDWW}\}.$

The lemma follows from the two observations given above. For example, any strategy σ_i which satisfies (2.4) is outcome-equivalent to one of the classes in $\tilde{\Sigma}_i$ by observation 1 above. For player $-i$'s strategies, by observation 1, we can focus on the class of strategies of the form Inab(C)b(D) for $b(C), b(D) \in \{T, E, W\}$. By observation 2, InCTE is outcome-equivalent to InCEE, and similarly for the other cases in which $b(C) = T$ or $b(D) = T$, concluding that focusing on $\tilde{\Sigma}_{-i}$ is without loss of generality.

Now we show that the strategies InDEW and InDWW are strictly dominated by a mixed strategy $(\text{Out}, \text{InDEE}; 1-p, p)$ and $(\text{Out}, \text{InDWE}; 1-p, p)$, respectively. Then the payoffs obtained by player i against player $-i$'s strategies InCb(C)b(D) are given in the matrices given in Tables 2.3 and 2.3.

Now we derive a condition under which $(\text{Out}, \text{InDEE}; 1-p, p)$ strictly dominates InDWE (the case of InDWW is the same). Note that as long as Out is played with a positive probability (i.e., $p < 1$) and the opponent plays InDb(C)b(D), $(\text{Out}, \text{InDEE}; 1-p, p)$ is strictly better than InDWE.[7] Thus we need the following conditions:

$$p\{(1-\delta)(1+g+\eta) - \varepsilon\} > (1-\delta)(1+g-\alpha) - \varepsilon \tag{2.5}$$

$$p\{(1-\delta)(1+g-\gamma) - \varepsilon\} > (1-\delta)(1+g-\beta) - \varepsilon \tag{2.6}$$

Consider the condition (2.5). If the left-hand side $(1-\delta)(1+g+\eta) - \varepsilon < 0$, then InDWE is strictly dominated by Out.[8] Thus let us assume $(1-\delta)(1+g+\eta) - \varepsilon > 0$. In this case, (2.5) becomes

$$p > \frac{(1-\delta)(1+g-\alpha) - \varepsilon}{(1-\delta)(1+g+\eta) - \varepsilon}.$$

Observe that the right-hand side is strictly less than 1.

[7] Note the assumption $\eta \leq \varepsilon$.

[8] In particular, if $(1-\delta)(1+g+\eta) - \varepsilon < 0$, then both InDWE and InDEE are strictly dominated by Out.

Table 2.3. Chapter 2 Proposition 5: First Round of Deletion against Out and InC

	Out	InCEE	InCEW	InCWE	InCWW
Out	0	0	0	0	0
InDEE	$-\varepsilon$	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$
InDWE	$-\varepsilon$	$(1-\delta)(1+g-\alpha)-\varepsilon$	$(1-\delta)(1+g-\beta)-\varepsilon$	$(1-\delta)(1+g-\alpha)-\varepsilon$	$(1-\delta)(1+g-\beta)-\varepsilon$
InDEW	$-\varepsilon$	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$
InDWW	$-\varepsilon$	$(1-\delta)(1+g-\alpha)-\varepsilon$	$(1-\delta)(1+g-\beta)-\varepsilon$	$(1-\delta)(1+g-\alpha)-\varepsilon$	$(1-\delta)(1+g-\beta)-\varepsilon$

Table 2.4. Chapter 2 Proposition 5: First Round of Deletion against InD

	InDEE	InDEW	InDWE	InDWW
Out	0	0	0	0
InDEE	$(1-\delta)\eta-\varepsilon$	$-(1-\delta)\gamma-\varepsilon$	$(1-\delta)\eta-\varepsilon$	$-(1-\delta)\gamma-\varepsilon$
InDWE	$(1-\delta)\eta-\varepsilon$	$-(1-\delta)\gamma-\varepsilon$	$(1-\delta)\eta-\varepsilon$	$-(1-\delta)\gamma-\varepsilon$
InDEW	$-(1-\alpha)-\varepsilon$	$-(1-\delta)\beta-\varepsilon$	$-(1-\delta)\alpha-\varepsilon$	$-(1-\delta)\gamma-\varepsilon$
InDWW	$-(1-\alpha)-\varepsilon$	$-(1-\delta)\beta-\varepsilon$	$-(1-\delta)\alpha-\varepsilon$	$-(1-\delta)\gamma-\varepsilon$

Now the condition for p is

$$\frac{(1-\delta)(1+g-\alpha)-\varepsilon}{(1-\delta)(1+g+\eta)-\varepsilon} < p < \frac{(1-\delta)(1+g-\beta)-\varepsilon}{(1-\delta)(1+g-\gamma)-\varepsilon} \tag{2.7}$$

Note that, since the right-hand side of (2.7) is less than 1, p can be taken to be less than 1 as long as p satisfying the above condition exists. Such p exists if and only if

$$\frac{(1-\delta)(1+g-\alpha)-\varepsilon}{(1-\delta)(1+g+\eta)-\varepsilon} < \frac{(1-\delta)(1+g-\beta)-\varepsilon}{(1-\delta)(1+g-\gamma)-\varepsilon}$$

This is equivalent to

$$\begin{aligned}\varepsilon &> (1-\delta)\frac{(\gamma+\alpha+\eta-\beta)(1+g)-\alpha\gamma-\beta\eta}{\gamma+\alpha+\eta-\beta} \\ &= (1-\delta)\left\{1+g-\frac{\alpha\gamma+\beta\eta}{\gamma+\alpha+\eta-\beta}\right\}\end{aligned}$$

The case of InDWW is analogous.

Note that Out, InDEE, and InDEW is not deleted in the first round since they are optimal against some CPS.[9]

Now consider a history $h = ((\mathrm{In}, \mathrm{In}), (D, C))$. The payoff matrix is given in Table 2.5.

The strategies InCEW and InCWW are conditionally dominated and are deleted. Therefore, the strategies σ such that $\sigma_i((\mathrm{In}, \mathrm{In}), (C, D)) = W$ are not extensive-form rationalizable. The deterrence strategy is included in the class of such strategies.

□

[9]Indeed, Out is optimal against μ such that $\mu(\sigma(h^0) = Out) = 1$ and the latter two are optimal against μ such that $\mu(\bar{\sigma}) = 1$ with

$$\begin{aligned}&\bar{\sigma}(h^0) = \mathrm{In} \\ &\bar{\sigma}((\mathrm{In}, \mathrm{In})) = C \\ &\bar{\sigma}((\mathrm{In}, \mathrm{In}), (D, C)) = E\end{aligned}$$

Table 2.5. Chapter 2 Proposition 5: Second Round

	InCEE	InCEW	InCWE	InCWW
InDEE	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$
InDEW	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$	$(1-\delta)(1+g+\eta)-\varepsilon$	$(1-\delta)(1+g-\gamma)-\varepsilon$

This shows that a "deviation" in period 1 will not be punished.

The previous result shows that perfectly cooperative outcome is not strategically stable in the sense of extensive-form rationalizability. The result, however, does not exclude the possibility of credibility of W in case of other histories. Making use of the existence of exit option, we can show that, irrespective of the outcome path, the outcome in which unilateral defection is followed by (W,W) is not strategically stable.

To show this result, we need additional condition for parameter values and epistemic restriction about the players' beliefs.[10] [11]

Proposition 6 If the conditions

$$\delta(1+g+\eta)-\eta<0 \tag{2.8}$$

or

$$\delta(1+g-\beta)-\eta<0 \quad \textbf{and} \quad \eta>\delta\frac{(\gamma+\alpha+\eta-\beta)(1+g)-\alpha\gamma-\beta\eta}{\gamma+\alpha+\eta-\beta} \tag{2.9}$$

are met, then outcomes of the form $z=((\mathrm{In},\mathrm{In}),\mathbf{a}^1,\mathbf{b}^1,\ldots,\mathbf{a}^{t-1},\mathbf{b}^{t-1},(D,C),(W,W))$ is not supported by strategy profiles which survive Δ-rationalizability with the belief restriction

$$\Delta_1=\Delta_2=\{\mu:\ \mu(\sigma(h)=T|h)\geq\zeta \text{ for } h=((\mathrm{In},\mathrm{In}),\mathbf{a}^1,\mathbf{b}^1,\ldots,\mathbf{a}^{t-1})\}$$

for some $\zeta>0$.[12]

Example 3 Again the first condition (2.10) is easy to satisfy. The parameter values $g=l=1$, $\delta=0.4$, $\varepsilon=0.5$, $\eta=0.4$, $\alpha=1$, $\beta=2$, and $\gamma=4$ satisfy the second condition (2.11) while violating the first.

[10] The same result applies if we allow elimination of weakly dominated strategies as the solution concept. To make the underlying epistemic restriction more transparent, here we use the belief restriction as in the main text.

[11] One caveat here is that Δ-rationalizability is not monotone in belief restrictions and thus a stronger epistemic restriction might broaden the set of rationalizable strategies. Still, the following result holds as long as the belief restriction does not exclude (i.e., does not put probability zero to) the outcome z.

[12] Δ_i thus defined is nonempty and closed, and thus is "regular" in the sense of Battigalli (2003) (Regularity in Battigalli's paper also imposes a restriction on the first-order restriction about the state of nature, which is trivially satisfied in the current setting). This ensures the existence of Δ-rationalizable strategy profiles.

Proof of Proposition 6 First, let $h = ((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \ldots, \mathbf{a}^{t-1})$ and

$$\Phi = -\varepsilon + (1-\delta)\sum_{\tau=1}^{t-2} \delta^{\tau-1}\{u^A(\mathbf{a}^\tau) + u^B(\mathbf{b}^\tau)\} + (1-\delta)\delta^{t-1}u^A(\mathbf{a}^{t-1})$$

be the discounted sum of payoffs up to history h

Denote by $Tab(C)b(D)$ an equivalence class of strategies such that

$$\begin{aligned}
&\sigma(h) = T \\
&\sigma(h, (T, T)) = a \\
&\sigma(h, (T, T), (a, C)) = b(C) \\
&\sigma(h, (T, T), (a, D)) = b(D),
\end{aligned}$$

by Exit the class of strategies such that $\sigma(h) = E$, and by War the class of strategies such that $\sigma(h) = W$. By the similar consideration as in the proof of Proposition 5, it suffices to focus on player i's strategies

$$\{Exit, TDEE, TDWE, TDEW, TDWW\}$$

and player $-i$'s strategies

$$\{Exit, War, TCEE, TCEW, TCWE, TCWW, TDEE, TDEW, TDWE, TDWW\}.$$

Now we show that the (equivalence class of) strategy TDEW is (conditionally) dominated by a mixed strategy $(\text{Exit}, \text{TDEE}; 1-p, p)$. The payoff matrices are given in Tables 2.6, 2.7, and 2.8.

Table 2.6. Chapter 2 Proposition 6: First Round against Exit and War

	Exit	War
Exit	$\Phi + (1-\delta)\eta$	$\Phi - (1-\delta)\gamma$
TDEE	$\Phi + (1-\delta)\eta$	$\Phi - (1-\delta)\gamma$
TDWE	$\Phi + (1-\delta)\eta$	$\Phi - (1-\delta)\gamma$
TDEW	$\Phi + (1-\delta)\eta$	$\Phi - (1-\delta)\gamma$
TDWW	$\Phi + (1-\delta)\eta$	$\Phi - (1-\delta)\gamma$

Table 2.7. Chapter 2 Proposition 6: First Round against TC

	TCEE	TCEW	TCWE	TCWW
Exit	$\Phi + (1-\delta)\eta$	$\Phi + (1-\delta)\eta$	$\Phi + (1-\delta)\eta$	$\Phi + (1-\delta)\eta$
TDEE	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$
TDWE	$\Phi + (1-\delta)\delta(1+g-\alpha)$	$\Phi + (1-\delta)\delta(1+g-\beta)$	$\Phi + (1-\delta)\delta(1+g-\alpha)$	$\Phi + (1-\delta)\delta(1+g-\beta)$
TDEW	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$
TDWW	$\Phi + (1-\delta)\delta(1+g-\alpha)$	$\Phi + (1-\delta)\delta(1+g-\beta)$	$\Phi + (1-\delta)\delta(1+g-\alpha)$	$\Phi + (1-\delta)\delta(1+g-\beta)$

Table 2.8. Chapter 2 Proposition 6: First Round against TD

	TDEE	TDEW	TDWE	TDWW
Exit	$\Phi + (1-\delta)\eta$	$\Phi + (1-\delta)\eta$	$\Phi + (1-\delta)\eta$	$\Phi + (1-\delta)\eta$
TDEE	$\Phi + (1-\delta)\delta\eta$	$\Phi + (1-\delta)\delta(-\gamma)$	$\Phi + (1-\delta)\delta\eta$	$\Phi + (1-\delta)\delta(-\gamma)$
TDWE	$\Phi + (1-\delta)\delta\eta$	$\Phi + (1-\delta)\delta(-\gamma)$	$\Phi + (1-\delta)\delta\eta$	$\Phi + (1-\delta)\delta(-\gamma)$
TDEW	$\Phi + (1-\delta)\delta(-\alpha)$	$\Phi + (1-\delta)\delta(-\beta)$	$\Phi + (1-\delta)\delta(-\alpha)$	$\Phi + (1-\delta)\delta(-\beta)$
TDWW	$\Phi + (1-\delta)\delta(-\alpha)$	$\Phi + (1-\delta)\delta(-\beta)$	$\Phi + (1-\delta)\delta(-\alpha)$	$\Phi + (1-\delta)\delta(-\beta)$

Due to the belief restriction Δ, there is no CPS against which TDEW is optimal if $(\text{Exit}, \text{TDEE}; 1-p, p)$ with $p < 1$ weakly dominates TDEW at the history h. Then the condition is as follows:

$$(1-p)\{\Phi + (1-\delta)\eta\} + p\{\Phi + (1-\delta)\delta(1+g+\eta)\} > \Phi + (1-\delta)\delta(1+g-\alpha) \tag{2.10}$$

$$(1-p)\{\Phi + (1-\delta)\eta\} + p\{\Phi + (1-\delta)\delta(1+g-\gamma)\} > \Phi + (1-\delta)\delta(1+g-\beta) \tag{2.11}$$

Arranging terms, (2.10) yields

$$p\{\delta(1+g+\eta) - \eta\} > \delta(1+g-\alpha) - \eta$$

If $\delta(1+g+\eta)-\eta < 0$, then TDWE is strictly dominated. Assume $\delta(1+g+\eta)-\eta > 0$. Then

$$p > \frac{\delta(1+g-\alpha)-\eta}{\delta(1+g+\eta)-\eta},$$

where the right-hand side is less than 1.

The second condition (2.11) is equivalent to

$$p\{\delta(1+g-\gamma) - \eta\} > \delta(1+g-\beta) - \eta$$

If $\delta(1+g-\gamma)-\eta \geq 0$, then the above condition is never satisfied. Assume $\delta(1+g-\gamma)-\eta < 0$. Then

$$p < \frac{\delta(1+g-\beta)-\eta}{\delta(1+g-\gamma)-\eta}.$$

The right-hand side must be nonnegative. Hence $\delta(1+g-\beta)-\eta < 0$ must hold, and also the right-hand side is less than 1.

Therefore we must have the condition

$$\frac{\delta(1+g-\alpha)-\eta}{\delta(1+g+\eta)-\eta} < \frac{\delta(1+g-\beta)-\eta}{\delta(1+g-\gamma)-\eta}$$

This is equivalent to

$$\eta > \delta \frac{(\gamma + \alpha + \eta - \beta)(1 + g) - \alpha\gamma - \beta\eta}{\gamma + \alpha + \eta - \beta}$$

A similar procedure applies for deleting TDWW.

Now consider the history $((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \ldots, \mathbf{a}^{t-1}, \mathbf{b}^{t-1}, (D, C))$. The payoff matrix is given in Table 2.9.

Thus the strategies TCEW and TCWW are conditionally dominated. Therefore, the set of strategies $\{\sigma : \sigma(((\text{In}, \text{In}), \mathbf{a}^1, \mathbf{b}^1, \ldots, \mathbf{a}^{t-1}, \mathbf{b}^{t-1}, (C, D))) = W\}$ is not Δ-rationalizable and the result follows. $\square$

Table 2.9. Chapter 2 Proposition 6: Second Round

	TCEE	TCEW	TCWE	TCWW
TDEE	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$
TDEW	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$	$\Phi + (1-\delta)\delta(1+g+\eta)$	$\Phi + (1-\delta)\delta(1+g-\gamma)$

Chapter 3 Credibility of Deterrence and Extensive-Form Rationalizability: Incomplete-Information Case

3.1 Introduction

Isogai (2016) and Isogai (2017) has explored the possibility that mutually destructive punishment is not always a rationalizable way of deterring a deviation in the existence of outside option. The gist is that through dynamic interaction, a certain choice of action conveys players' intention by the fact that a particular response is inferior to others. If we impose an assumption that players are rational and believes that the opponents are rational as much as possible, then the inferior response would not be played by rational players (forward induction). The results derived in the previous papers show that, as long as the parameter restrictions (and the epistemic assumption of rationality and common strong belief of rationality) are met, mutually destructive punishment cannot be used as a credible deterrence of deviation.

The previous results not only have implications to the reality which explains the failure of punishment observed in real cartel cases but also show the implication of rationality and common strong belief of rationality (RCSBR) assumption (Battigalli and Friedenberg, 2012) that underlies extensive-form rationalizability. Then a natural question would arise: is there a setting under which punishment is compatible with RCSBR? This paper studies that possibility under incomplete information. To show the pervasiveness of the previous results on rationalizability, we mainly consider a model in which there is no payoff uncertainty about each stage game (except for the section which shows how the relaxation of the assumption works).

To that objective, we would like to fix the parameter values in which punishment option is not rationalizable between two players, and instead introduce the multiplicity of players so that an occurrence of punishment, which is observable to the other players, might serve as a signal of the player's intention to be a "tough" punisher. To focus on the effect of observed punishment to other players who plays the game in the future, I consider a model of short-lived players, who are sequentially matched with a single long-lived player.

As I illustrate in the next section, under a complete information environment, "backward-induction" argument prevails, and the non-rationalizability result is not overturned by finitely many repetition of the stage games. Therefore, signaling of incomplete information is crucial; I assume that the number of short-lived players is randomly chosen. The long-lived player knows the realization of the number of short-lived players while the short-lived players cannot observe it. The stage game played by the long-lived player and each short-lived player is understood as a prisoners' dilemma game followed by punishment choice, which we simplify to sequential choice by each player: first the short-lived player chooses whether to "cooperate" or "defect," after which the long-lived player chooses whether to "punish" or "tolerate" the short-lived player. While the short-lived players observe the past occurrences of punishment, they do not observe whether the short-lived players have defected or not. Our focus in the paper is then whether a punishment occurs in equilibrium and whether such strategy is rationalizable.

Naively, we might think that when there are sufficiently many short-lived players, then the long-lived player can punish a deviator and discipline the short-lived players while the short-lived players infer the "stock" of the short-lived players and thus obey the long-lived player. However, it turns out that this is not the case in our model. There is no equilibrium in which the long-lived player uses a class of punishment strategies with which the long-lived player punishes a deviant short-lived player as long as there remain sufficiently many short-lived players. The reason is as follows: since the short-lived players are symmetric, their actions in the absence of an occurrence of punishment are the same. Then the possible outcomes would be either one in which all short-lived players cooperate or one in which all defect. The former outcome is not supported in an equilibrium since a short-lived player's unilateral deviation would not be punished. In the latter case, the only possibility is that an occurrence of punishment might change the short-lived players' responses. It turns out, however, that the number of short-lived players who might change their responses is fewer than the number which rationalizes the long-lived player's punishment choice.

This problem arises from the informational assumption that failures to punishment are not observed and the common knowledge assumption of payoffs. In the sequel, I show that the existence of equilibria with punishments is recovered when we modify these assumptions.

In addition to those equilibrium analyses, I also provide results on rationalizability. Since the games I consider are incomplete-information games, the appropriate solution concept is Δ-rationalizability formalized by Battigalli and Siniscalchi (2003). I show that, except for the case in which punishment is obviously dominated, strategies which involve punishment are rationalizable. This is in stark contrast with the complete-information cases in which there is no rationalizable equilibrium (i.e., equilibrium consisting of rationalizable strategies) which involves punishment. My results in this thesis show that incomplete-information models might give punishment the signaling role of private information, which makes the action rationalizable.

At the same time, without additional belief restrictions, rationalizability does not have enough predictive power of possible outcomes of the game. In the Appendix, I suggest an approach of using belief restrictions to obtain a result which supports outcomes in which an occurrence of punishment can deter future defections. While the result is extremely simple, the approach shows a possible explanation of player's incentive to use punishment, which is difficult to capture using the equilibrium concept.

The current study is important in the sense that it studies how the previous result on rationalizability of punishment can be overturned by a modification of the model. Furthermore, it also shows that the signaling effect of observable punishment still depends critically on the observability of failure of punishment, rather than its observability.

It also sheds light on the possibility of amnesty program which might be exploited by a multiproduct firm which can signal its intention to punish other cartel firms. While the results in the previous studies suggest that the only rationalizable equilibrium outcome is the one in which a leniency application does not occur under some circumstance, the current result shows that if there is an incomplete information, the leniency application is rationalizable. The precise argument is discussed in Isogai and Shen (2017).

The paper is structured as follows: in section 2, I provide an illustrative example which shows that finitely many repetition of games does not help recover rationalizability of punishment. In section 3, I define our baseline random-player game and show that a natural class of punishment strategies cannot constitute Perfect Bayesian equilibria. In sections 4, I illustrate that introduction of payoff uncertainty gives rise to an equilibrium in which an occurrence of punishment deters future defection. The section 5 discusses the relationship between the current paper and other studies using rationalizability. The Appendix A provides how to formally construct the Harsanyi type space for the random-player game, Appendix B considers a variant of the main model in which punishment can be used in equilibrium, and Appendix C provides a brief discussion about Δ-rationalizability approach. Since the current paper involves several versions of the model, the notations are somewhat involved. Appendix D summarizes the notations used in each section.

3.2 Non-Rationalizability Result Revisited: Finitely Many Games Does Not Help

To motivate the random-player game discussed in the next section, I provide an example which illustrates why complete information of the number of short-lived players prohibit the punishment from being credible. Consider the following two-period finitely repeated game, whose stage game at $t = 1, 2$ consists of the following three phases:

1. At the first phase, the long-lived player (player 0) and the short-lived player (player $t = 1, 2$) decides whether to enter the game in subsequent phases. The entry cost is denoted by $\varepsilon > 0$.

2. Only if both of the players decide to enter, the players are involved in the

prisoners' dilemma game:

Table 3.1. Chapter 3: First Stage of the Finite Case

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

where $g, l > 0$.

3. After observing the outcome in the second phase, players play the following coordination game:

Table 3.2. Chapter 3: Second Stage of the Finite Case

	T	W
T	$0, 0$	$-\gamma, -\alpha$
W	$-\alpha, -\gamma$	$-\beta, -\beta$

where $\gamma > \beta > \alpha > 0$. Note that under the parameter assumptions, (T, T) and (W, W) are the stage-game (pure strategy) Nash equilibria.

The game is simply the twice repeated game version of the model in Isogai (2017a).

The result in the paper can be applied without a major change:

Theorem 2 If

$$0 < \varepsilon < 1$$
$$(1 + g - \varepsilon)(\beta - \alpha) > (1 + g - \alpha - \varepsilon)$$

are satisfied, then there is no subgame-perfect equilibrium which consists of extensive-form rationalizable strategies under which punishment occurs.

<u>Sketch of Theorem 2</u> As in the analysis in Isogai (2017a), in the first stage, the short-lived players strategy "enter, play D, and chooses W if the outcome in the prisoners' dilemma phase was (D, C)" is deleted. Then in the second stage, the long-lived player's strategy which chooses W after observing (C, D) in the second period is deleted. In the third stage, the long-lived player's strategy which chooses W after observing (C, D) in the first period is deleted. □

Intuitively, the "backward induction" holds for extensive-form rationalizable equilibrium since the short-lived players' rationalizable strategies are not history-dependent (i.e., the same through time).

With this result in mind, from the next section on, I consider a model which studies under what conditions the punishment (W) can be played in an equilibrium which is robust to forward-induction reasoning. To simplify the analysis, the stage game in the following sections are reduced form; i.e., the short-lived players only choose whether to cooperate or not and do not have a choice to punish the long-lived player since short-lived players' punishment choice is in any case not rationalizable.[1]

3.3 Random-Player Model

3.3.1 Basic Structure

Next I consider a model of random-player game (with unbounded support) in which one long-lived player interacts with multiple short-lived players. The game falls in random-player games as in Myerson (1998) and Milchtaich (2004). I assume that the set of short-lived players are randomly chosen according to the probability distribution $d(n) := (1 - \xi)\xi^{n-1}$ for $n = 1, 2, \ldots$ with $\xi \in (0, 1)$.[2] For a realization n, let $\{1, 2, \ldots, n\}$ be the set of short-lived players. Denote the long-lived player by 0. In the following, when we denote the players including the long-lived player, we use the index $i = 0, 1, 2, \ldots$ while we use $s = 1, 2, \ldots$ when we only mention short-lived players.

The timeline of the game is as follows:

Period 0 The number of short-lived players n is chosen according to the distribution $d(n)$. Only the long-lived player observes the realization n.

Period $s = 1, \ldots, n$ The short-lived player, denoted by s, is chosen and plays the

[1]This simplification is consistent with the more "complete" model, which allows short-lived players to choose punishment because extensive-form rationalizability is order-independent up to outcome-equivalence (Chen and Micali, 2013).

[2]We can construct a Harsanyi type space. The detail is in the Appendix.

game with the long-lived player. The game consists of the following two stages:

1. In the first stage, the short-lived player chooses between C and D, where the stage-game payoffs for long-lived and short-lived players are

$$u^1(C) = (u_0^1(C), u_s^1(C)) = (1, 1)$$
$$u^1(D) = (u_0^1(D), u_s^1(D)) = (-l, 1 + g)$$

   where $g, l > 0$. The interpretation is that in a prisoners' dilemma game, the long-lived player commits to play C and the short-lived player decides to choose C or D.[3]

2. In the second stage, the long-lived player chooses between T and W, where the stage-game payoffs are

$$u^2(T) = (u_0^2(T), u_s^2(T)) = (0, 0)$$
$$u^2(W) = (u_0^2(W), u_s^2(W)) = (-\alpha, -\gamma)$$

   where $\alpha, \gamma > 0$. The action T means tolerance and W means punishment (or trigger of War). Note that the choice W is costly for the long-lived player.

Note that the payoff functions for short-lived players are same across s.

Throughout we assume that the punishment is strong enough to deter a defection (D) when it occurs with certainty: $g < \gamma$.

The long-lived player knows the whole history of the game. As to the short-lived players' information, first we consider a case in which the short-lived players can only observe the past occurrence of punishment. The formal specification of information structure is given in the next section.

The payoff for the long-lived player is the sum of payoffs in each periods[4] and the payoff of the short-lived player is the within-period payoffs. I specify the payoff

[3]Assuming instead that both players play the prisoners' dilemma game and then playing the punishment-stage game as in Isogai (2017a) does not change the qualitative implication of the model.

[4]Since n is finite with probability one, the (interim) payoff is finite with probability one. Moreover, since the payoffs from each period is bounded and the distribution $d(n)$ is geometric, ex-ante payoff is also bounded.

in the dynamic game in the next section, where I define the set of histories and outcomes.

3.3.2 Histories, Payoff Functions, Strategies, and Solution Concepts

The set of long-lived player's (non-terminal) histories consists of the following:

$$\mathcal{H}_0^0(n) := \{(n) : \ n = 1, 2, \ldots\}$$
$$\mathcal{H}_0^t(n) = \{h = (n, a^1, b^1, a^2, \ldots, a^t) : \ n \in \mathbb{N}, a^\tau \in \{C, D\}, \ b^\tau \in \{T, W\}\}$$

Let $\mathcal{H}(n) := \cup_{0 \le t \le n} \mathcal{H}_0^t(n)$ and $\mathcal{H}_0 := \cup_{n \ge 1} \mathcal{H}_0(n)$.

The set of short-lived players' histories is

$$\mathcal{H}_s := \{(b^1, b^2, \ldots, b^{t-1}) : \ b^\tau \in \{T, W\}\}$$

The set of outcomes of the game is given as follows:

$$\mathcal{Z}_n := \{(n, a^1, b^1, a^2, \ldots, a^n, b^n) : \ a^\tau \in \{C, D\}; \ b^\tau \in \{T, W\}\} \quad \text{for } n \in \mathbb{N}$$
$$\mathcal{Z} := \bigcup_{n \in \mathbb{N}} \mathcal{Z}_n$$

The payoff functions are

$$\tilde{U}_0(z_n) = \sum_{t=1}^{n} \{u_0^1(a^t) + u_0^2(b^t)\}$$
$$\tilde{U}_s(z_n) = u_s^1(a^s) + u_s^2(b^s) \quad \text{for } s \le n$$

for $z_n \in \mathcal{Z}_n$.

A (pure) strategy for the long-lived player is a mapping $\sigma_0 : \mathcal{H}_0 \to \{T, W\}$ and a strategy for short-lived player s is a mapping $\sigma_s : \mathcal{H}_s \to \{C, D\}$. The set of player i's strategies is denoted by Σ_i and $\Sigma = \prod_i \Sigma_i$. Note that $\Sigma_0 = \cup_{n \ge 1} \{T, W\}^{\mathcal{H}(n)}$ and $\Sigma_s = \{C, D\}^{\mathcal{H}_s}$, and thus the sets of strategies are at most countable. Player i's mixed strategy is a probability distribution over Σ_i. Let $\zeta : \mathbb{N} \times \Sigma_0 \times (\prod_{s \in \mathbb{N}} \Sigma_s) \to \mathcal{Z}$ be the outcome function induced by the number of players and the players' strategies.[5] The strategic-form payoff function U_i of player i is given by

$$U_i(n, \sigma) = \tilde{U}_i(\zeta(n, \sigma)).$$

We say that the long-lived player's strategy σ_0 is consistent with history h_0 if there exist opponents' strategies $(\sigma_s)_s$ such that $\zeta(n, \sigma_0, (\sigma_s)_s)$ goes through h_0 (with positive probability). Similarly, we say that a short-lived player's strategy σ_s is consistent with history h_s if there exist a number $n \geq 1$ and opponents' strategies (σ_0, σ_{-s}) such that $\zeta(n, \sigma_s, \sigma_0, \sigma_{-s})$ goes through h_s (with positive probability).[6] The set of player i's strategies consistent with history h_i is denoted by $\Sigma_i(h_i)$.

The players' beliefs are given by conditional probability systems, which I define below:

Definition 8 (Conditional Probability System (Rényi, 1955)) A conditional probability system (CPS) of the long-lived player 0 is a mapping μ_0 which maps long-lived player's history $h_0 \in \mathcal{H}_0$ to a belief (i.e., probability distribution) over short-lived players' strategies $\mu_0((\sigma_s)_s|h_0)$ which satisfies

Properness $\mu_0(\Sigma_{-i}(h)|h) = 1$ for any $h \in \mathcal{H}_0$.

Consistency For h and h', $\mu_0((\sigma_s)_s|h) = \mu_0((\sigma_s)_s|h')\mu_i(\Sigma_{-i}(h')|h)$.

Similarly, a CPS of short-lived player s is a mapping μ_s which maps the short-lived player s's history $h_s \in \mathcal{H}_s$ to a belief over the number of short-lived players and long-lived player's strategies $\mu_s(n, \sigma_0, (\sigma_{s'})_{s' \neq s}|h_s)$ with the same restrictions as above.

Denote by $\mathcal{C}_i$ the set of CPSs for player i.

One solution concept we use is Perfect Bayesian equilibrium (Fudenberg and Tirole, 1991). We give the definition of the equilibrium in our current setting:

Definition 9 (Perfect Bayesian Equilibrium) A strategy profile $(\sigma_i)_{i=0,1,2,\ldots}$ constitutes a Perfect Bayesian equilibrium (PBE) if there exist CPSs $(\mu_i)_i$ such that

[5] When it involves a mixed strategy, the outcome function is a probability distribution over $\mathcal{Z}$.

[6] More precisely, the condition is that the appropriate projection of $\zeta(n, \sigma_s, \sigma_0, \sigma_{-s})$ goes through h_s with positive probability.

1. The players are certain of the distribution of the number of short-lived players:

$$\sum_{\sigma_{-s}} \mu_s(n, \sigma_{-s}|\emptyset) = d(n)$$

2. The players are certain of the opponents' strategies: for any history h_i,

$$\sum_{n} \mu_i(n, \sigma|h_i) = 1.$$

3. At each history $h_i \in \mathcal{H}_i$, the action $\sigma_i(h_i)$ maximizes players i's expected payoff evaluated by the belief $\mu_i(\cdot|h_i)$.

As a benchmark, denote by $(\sigma_0^{NP}, \sigma_s^{NP})$ the profile of "non-punishment strategy profile" which is defined as

$$\sigma_0^{NP}(h_0) = T$$
$$\sigma_s^{NP}(h_s) = D$$

for any histories h_0 and h_s. Note that non-punishment strategy profile is always a PBE.

Another solution concept is Battigalli and Siniscalchi (2003)'s Δ-rationalizability. While the solution concept does not necessarily assume common certainty of strategies, it captures the idea of forward induction. One requirement of Δ-rationalizability is that the players be sequentially rational:

Definition 10 (Sequential Rationality) Player 0's strategy σ_0 is called sequentially rational against $\mu_0 \in \mathcal{C}_0$ if and only if at any history $h \in \mathcal{H}(n)$ for which $\sigma_0 \in \Sigma_0(h)$,

$$\sum_{\sigma_{-0}} U_0(n, \sigma_0, \sigma_{-0})\mu_0(\sigma_{-0}|h) \geq \sum_{\sigma_{-0}} U_0(n, \sigma_0', \sigma_{-0})\mu_0(\sigma_{-0}|h)$$

for any $\sigma_0' \in \Sigma_0(h)$. Similarly, short-lived player s's strategy σ_s is called sequentially rational against $\mu_0 \in \mathcal{C}_s$ if and only if at any history h for which $\sigma_s \in \Sigma_s(h)$,

$$\sum_{n,\sigma_{-s}} U_s(n, \sigma_s, \sigma_{-s}\mu_s(n, \sigma_{-s}|h) \geq \sum_{n,\sigma_{-s}} U_s(n, \sigma_s', \sigma_{-s})\mu_s(n, \sigma_{-s}|h)$$

for any $\sigma'_s \in \Sigma_s(h)$.

The difference from the optimality requirement in PBE is that the sequential rationality concept requires the maximization of continuation payoffs only after histories for which the strategy is consistent with.

In Δ-rationalizability, we can also impose additional restrictions on the players' possible beliefs. The restrictions are given by $\Delta = (\Delta_i)_i \subset (\mathcal{C}_i)_i$ of beliefs about the opponent's strategies.

Definition 11 (Δ-rationalizability) Let Δ be given. Consider the following procedure:

1. Let $\Sigma_i(\Delta, 0) := \Sigma_i$
2. For $t \geq 0$ and $\Sigma(\Delta, t) = \Sigma_i(\Delta, t) \times \Sigma_{-i}(\Delta, t)$ given, $\sigma_i \in \Sigma_i(\Delta, t+1)$ if and only if there exists a $\mu_i \in \Delta_i$ such that
   (a) σ_i is sequential rational against μ_i
   (b) If at a history h, $\Sigma_{-i}(h) \cap \Sigma_{-i}(\Delta, t) \neq \emptyset$ holds, then μ_i must satisfy $\mu_i(\Sigma_{-i}(\Delta, t)|h) = 1$.[7]

Then we say that $\Sigma(\Delta) := \cap_{t \geq 0} \Sigma(\Delta, t)$ is the set of Δ-rationalizable strategies.

Denote by $\bar{\Delta}$ the set of possible beliefs when there is a common certainty of ex-ante distribution $d(n)$ of the number of players:

$$\bar{\Delta} = \mathcal{C}_0 \times \left\{ \mu_s : \sum_{\sigma_{-s}} \mu_s(n, \sigma_{-s}|\emptyset) = d(n) \right\}^{\mathbb{N}}.$$

3.3.3 Equilibrium Analysis

In the current setting, it might seem that since the occurrence of punishment signals the "stock" of short-lived players that the long-lived player can threaten to cooperate, there should exist an equilibrium in which punishment occurs (and so

[7] Battigalli and Friedenberg (2012) uses a stronger condition for their Directed-rationalizability concept, which is the collection of Δ-rationalizable strategies across different belief restrictions. Our analysis below does not change with replacement of the conditions.

that C is played on some equilibrium path). If we focus on a class of long-lived player's strategies, this is not the case.

We consider the following class of strategies by the long-lived player. Given a positive integer I, construct the state variable $S_t^I(h^t)$ such that

$$S_1^I(h^1) = 1$$

$$S_{t+1}^I(h^{t+1}) = \begin{cases} \max\{S_t^I(h^t), t+1\} & \text{if } b_t = T \\ S_t^I(h^t) + I & \text{if } b_t = W \end{cases}$$

The state variable S_t^I denotes the sum of current period and the number of short-lived players who are "affected by the occurrence of punishment." If a punishment does not occur in a given period, then the number of affected short-lived players is unchanged and the number of short-lived players who has exited increases by 1. When a punishment occurs additionally, then it signals that there are I more short-lived players. Then consider the long-lived player's strategy that stipulates at history $h = (n, a^1, b^1, \dots, a^t)$,

- Play T if $a^t = C$.
- If $a^t = D$,
  - Choose W if $S_t^I(h^t) + I \leq n$.
  - Choose T otherwise.

Denote the strategy considered above by σ_0^I and call it "I-punishment" strategy.

An I-punishment strategy decides whether to punish a deviant short-lived player in terms of whether additional punishment pays to the cost of triggering a punishment. If short-lived players follow symmetric strategies (i.e., $\sigma_s = \sigma_{s'}$ for s, s'), then this is a natural class of strategies.

Let $\bar{I}$ be the integer that satisfies $\bar{I}(1+l) \geq \alpha > (\bar{I}-1)(1+l)$. That is, $\bar{I}$ is the smallest number of remaining short-lived players to justify the use of punishment.

The next theorem shows that there is no PBE in which I-punishment strategy is used against symmetric strategies taken by short-lived players.

Theorem 3 Assume $g \neq \gamma\xi^{\bar{I}}$, there is no Perfect Bayesian equilibrium (PBE) in which the long-lived player uses an I-punishment strategy and in which the short-lived players follow symmetric strategies.

Proof of Theorem 3 First, note that the outcome $(n, C, T, C, T, C, T, \ldots, C, T)$ is never supported by a PBE for any realization n. Suppose to the contrary that there exists such an equilibrium. Consider a history $h = (n, D)$. If the long-lived player chooses T, then she obtains the payoff $(n-1)(1+l)$ while she only obtains the payoff $(n-1)(1+l) - \alpha$ if she chooses W. Thus the long-lived player's best response is T, against which the short-lived players can profitably deviate to D. This is a contradiction.

Next, first assume that the short-lived players follow symmetric pure strategies. Suppose that there exists a PBE in which the long-lived player follows $\sigma_0^{\bar{I}}$. There are two exhaustive cases depending on the parameter values:

1. $g < \gamma\xi^{\bar{I}}$
2. $g > \gamma\xi^{\bar{I}}$

For the first case, the short-lived players' best response is to play C. However, then the resulting path is $(n, C, T, C, T, C, T, \ldots, C, T)$, which cannot be an equilibrium outcome. For the second case, consider a history (n, D, W) for $n \geq \bar{I} + 1$. Since the long-lived player follows $\sigma_0^{\bar{I}}$, it is common knowledge among the short-lived firms that there are at least $\bar{I} + 1$ short-lived players. Now consider the short-lived player $\bar{I} + 1$. His belief about the probability that there are $\bar{I}$ more short-lived players is $\xi^{\bar{I}}$. Then, by $g > \gamma\xi^{\bar{I}}$, his best-response is to play D. This means that the number of short-lived players who plays C is at most $\bar{I} - 1$, contradicting to the optimality of $\sigma_0^{\bar{I}}$.

For other $I > \bar{I}$, σ_0^I is not a long-lived player's equilibrium strategy since it is weakly dominated by $\sigma_0^{\bar{I}}$ (consider, for example, a history $h = (n, D)$ for which $n = \bar{I} + 1$).

For mixed strategy, note that the short-lived players must be indifferent between C and D. If long-lived player follows σ^I and $h = (n, D, W)$, then short-lived player $s(\geq 2)$'s belief about the probability that he is punished is ξ^{s-1}. Since the expected loss from being punished is $\gamma\xi^{s-1}$, which is decreasing in s, and since the long-lived player's strategy σ^I must be optimal, it must be the case that

$$\gamma\xi^1 > \gamma\xi^2 > \cdots > \gamma\xi^I \geq g,$$

where g is the benefit from playing D at the risk of being punished. The case

$\gamma\xi^I > g$ reduces to the pure-strategy case and the case $\gamma\xi^I = g$ is what we have excluded. □

It may be seen that the fact that punishment might perfectly deter defection collapses the equilibrium since a unilateral deviation by a short-lived player will be left unpunished. This is not the case; for example, if we introduce the possibility of "mutant" short-lived player who would always choose D, in equilibrium, the behaviors of other "strategic" short-lived players would be unchanged in equilibrium. More precisely, if the long-lived player would punish a short-lived player who plays D whenever there are "enough" short-lived players, then only the strategic short-lived players would play C in equilibriu. Then the long-lived player knows that a player who plays D is a mutant and thus will not choose W. Thus, by the same token, modification of the information structure to the case in which the occurrence of punishment is imperfectly observed does not help.

The problem here is thus one of the following:

1. Unobservability of "failure to punish,"
2. Absence of payoff uncertainty in stage game, and
3. The (Nash-)equilibrium assumption, i.e., the assumption that the players must rationally expect the other players' strategies.

The first and the second point is discussed in Appendix B and Section 4, respectively.

For the third point, under the equilibrium assumption, the short-lived players know that if the punishment is credible, then the other short-lived players, if any, would certainly cooperate with the long-lived player. Then, the long-lived player would have no incentive to punish a unilateral deviation by (only) one short-lived player.

Here I provide the following result which shows that Δ-rationalizability with unrestricted beliefs (except for the ex-ante probability of the number of short-lived players) gives only trivial set of strategies. A possible way of strengthening the predictive power of rationalizability is suggested in Appendix C.

For exposition, let

$$\tilde{\Sigma_0} = \{\sigma_0 : \ \sigma_0(h_0) = W \text{ if } h_0 \in \mathcal{H}^t(n) \text{ and } t + \bar{I} > n\}$$

be the set of long-lived player's strategies which stipulates punishment when there are not enough short-lived players.

Theorem 4 Consider the baseline random-player game with $\Delta = \bar{\Delta}$. Let $\bar{I}$ be as given above. If $g \leq \gamma\xi^{\bar{I}}$, then the sets of Δ-rationalizable strategies for the players are

$$\Sigma_0(\Delta) = \Sigma_0 \setminus \tilde{\Sigma}_0$$
$$\Sigma_s(\Delta) = \Sigma_s \text{ for } s = 1, 2, \ldots.$$

If $g > \gamma\xi^{\bar{I}}$, then the sets of Δ-rationalizable strategies for the players are

$$\Sigma_0(\Delta) = \{\mu_0 : \ \mu_0(h) = T \text{ for any } h\}$$
$$\Sigma_s(\Delta) = \{\mu_s : \ \mu_s(h) = D \text{ for any } h\} \text{ for } s = 1, 2, \ldots.$$

Proof of Theorem 4 First suppose $g \leq \gamma\xi^{\bar{I}}$. In the first round, the long-lived player's strategy $\sigma_0 \in \tilde{\Sigma}_0$ is deleted since it is strictly dominated. The other strategies are not deleted since there exists a belief for which the strategy is sequentially rational. In particular, σ_0 such that $\sigma_0(h(n)) = W$ is optimal against a belief μ_0 if $h(n) \in \mathcal{H}^t(n)$, $t + \bar{I} < n$, and μ_0 puts probability 1 to the short-lived players' strategies which chooses C if and only if the long-lived player chooses W at history $h(n)$. Hence $\Sigma_0(\Delta, 1) = \Sigma_0 \setminus \tilde{\Sigma}_0$. The short-lived players' strategies are not deleted: at any history h_s, player s's strategy σ_s is sequentially rational against some CPS μ_s because

- if $\sigma_s(h_s) = D$, let $\mu_s(n, \sigma_0^{NP}) = 1$ for some $n > s$ and non-punishment strategy σ_0^{NP}.
- if $\sigma_s(h_s) = C$, let $\sum_n \mu_s(n, \sigma_0^{\bar{I}}) = 1$; note the assumption $g \leq \gamma\xi^{\bar{I}}$.

Therefore, $\Sigma_s(\Delta, 1) = \Sigma_s$

In the second round, none of the long-lived player's strategy is deleted since the short-lived players' strategies are not deleted in the first round. By the same beliefs as above, the short-lived players' strategies are also not deleted. Hence the deletion procedure ends at the second round and the result follows.

Next suppose $g > \gamma \xi^{\bar{I}}$. In the first round, similar to the previous case, the long-lived player's strategy $\sigma_0 \in \tilde{\Sigma}_0$ is deleted: $\Sigma_0(\Delta, 1) = \Sigma_0 \setminus \tilde{\Sigma}_0$. In the second round, because now $\Sigma_0(\Delta, 1) = \Sigma_0 \setminus \tilde{\Sigma}_0$, short-lived players are certain that the probability that the long-lived player chooses W against D is at most $\xi^{\bar{I}}$. Thus by assumption, σ_s such that $\sigma_s(h) = C$ is never sequentially rational and thus $\Sigma_s(\Delta, 2) = \{\mu_s : \ \mu_s(h) = D \text{ for any } h\}$. In the third round, since $\Sigma_s(\Delta, 2) = \{\mu_s : \mu_s(h) = D \text{ for any } h\}$, σ_0 such that $\sigma_0(h) = W$ is never sequentially rational. Thus $\Sigma_0(\Delta, 3) = \{\mu_0 : \ \mu_0(h) = T \text{ for any } h\}$. The procedure ends in the forth round and the result follows. □

3.4 Model with Payoff Uncertainty

In this section, we illustrate that, under payoff uncertainty, there is an equilibrium in which punishment deters deviation.

We consider the same stage game as in the random-player game and consider only two short-lived player case (i.e., $n = 2$ is common knowledge). Instead, we introduce "commitment type," who incurs sufficiently large payoff loss by tolerating a deviation, i.e., choosing T after the short-lived player chose D in the same period. The extensive-form payoff functions are given below. Let $\bar{\mu} \in (0, 1)$ be the ex-ante probability of the long-lived player being the commitment type. With the complementary probability, the long-lived player is of the normal type, whose payoff specification is the same as before. We assume that $g > \bar{\mu}\gamma$, i.e., the probability of commitment type is not high enough to deter the short-lived players from defection if the normal type never punishes deviant short-lived players.

Formally, let $\mathcal{T} = \{\tau^n, \tau^c\}$ be the set of long-lived player's types, where τ^n (τ^c) denotes the normal (commitment) type. The sets of period-t ($t = 1, 2$) histories for the long-lived player are

$$\hat{\mathcal{H}}_0^1 = \mathcal{T} \times \{C, D\}$$
$$\hat{\mathcal{H}}_0^2 = \mathcal{T} \times \{C, D\} \times \{T, W\} \times \{C, D\}$$

The sets of histories for the short-lived players $t = 1, 2$ are

$$\hat{\mathcal{H}}_1 = \{\emptyset\}$$
$$\hat{\mathcal{H}}_2 = \{T, W\}$$

The set of outcomes is

$$\hat{\mathcal{Z}} := \mathcal{T} \times \{C, D\} \times \{T, W\} \times \{C, D\} \times \{T, W\}.$$

Strategies are defined as mappings from histories to actions.

The (extensive-form) payoff functions are now mappings $\hat{U}_0$ and $(\hat{U}_t)_{t=1,2}$ from $\hat{\mathcal{Z}}$ to $\mathbb{R}$ and are given by

$$\hat{U}_0(\tau, a^1, b^1, a^2, b^2) = \begin{cases} \sum_{t=1}^{2} \{u_0^1(a^t) + u_0^2(b^t)\} & \text{if } \tau = \tau^n \\ \sum_{t=1}^{2} \hat{U}_0^c(a^t, b^t) & \text{if } \tau = \tau^c \end{cases}$$

where

$$\hat{U}_0^c(a^t, b^t) = \begin{cases} -K & \text{if } a^t = D \text{ and } b^t = T \\ u_0^1(a^t) + u_0^2(b^t) & \text{otherwise} \end{cases}$$

for $K > 2(-l - \gamma)$, and

$$\hat{U}_t(\tau, a^1, b^1, a^2, b^2) = u_t^1(a^t) + u_t^2(b^t)$$

Now we show the existence of a Perfect Bayesian equilibrium in which cooperation by a short-lived player occurs with a positive probability because of the credible punishment. We assume that the long-lived player's gain from short-lived player's cooperation is large enough to make up for the cost of punishment: $\alpha < 1 + l$. Denote the strategy profile given below by $(\hat{\sigma}_i)_{i=0,1,2}$:

Long-lived player If the long-lived player is of normal type,

$$\hat{\sigma}_0(\tau^n, C) = T$$

$$\hat{\sigma}_0(\tau^n, D) = \begin{cases} T & \text{with probability } 1-\eta \\ W & \text{with probability } \eta \end{cases}$$

$$\hat{\sigma}_0(\tau^n, a^1, b^1, a^2) = T \quad \text{for } a^1, a^2 \in \{C, D\},\ b^1 \in \{T, W\}$$

for $\eta \in (0,1)$. If the long-lived player is of commitment type,

$$\hat{\sigma}_0(\tau^c, C) = T$$
$$\hat{\sigma}_0(\tau^c, D) = W$$
$$\hat{\sigma}_0(\tau^c, a^1, b^1, C) = T \quad \text{for } a^1 \in \{C, D\},\ b^1 \in \{T, W\}$$
$$\hat{\sigma}_0(\tau^c, a^1, b^1, D) = W \quad \text{for } a^1 \in \{C, D\},\ b^1 \in \{T, W\}$$

Short-lived players [8]

$$\hat{\sigma}_1(\emptyset) = D$$
$$\hat{\sigma}_2(T) = D$$
$$\hat{\sigma}_2(W) = \begin{cases} C & \text{with probability } \chi \\ D & \text{with probability } 1-\chi \end{cases}$$

for $\chi \in (0,1)$.

The mixing probabilities η and χ are derived in the proof of theorem.

Theorem 5 Under the assumption

$$1 \geq \bar{\mu}\frac{\gamma}{g}(1 + \gamma - g),$$

the strategy profile $(\hat{\sigma}_i)_{i=0,1,2}$ with an appropriate belief constitutes a Perfect Bayesian equilibrium. [9]

Proof of Theorem 5

[8] There is another equilibrium in which short-lived player 1's equilibrium strategy is C. In that equilibrium, punishment will not occur at period 1 on equilibrium path and thus the short-lived player 2 will deviate if the punishment did not occur.

[9] If the condition holds with equality, then a strategy profile in which the short-lived player 1's strategy replaced with arbitrary mixed strategy is also a PBE.

Long-lived player First consider the history $h_0^1 = (\tau^n, D)$. Since the normal-type long-lived player is indifferent between T and W, the indifference condition $\alpha = \chi(1+l)$ must be satisfied, so that $\chi = \frac{\alpha}{1+l}$. The optimality for the other histories (including those for commitment type) is obvious.

Short-lived player 2 After the history $h_2 = (W)$, Bayes' rule implies that the short-lived player 2's belief is

$$\frac{\bar{\mu}}{\bar{\mu} + (1-\bar{\mu})\eta}$$

Thus the indifference condition for short-lived player 2 is

$$\frac{\bar{\mu}}{\bar{\mu} + (1-\bar{\mu})\eta}\gamma = g$$
$$\iff \eta = \frac{\bar{\mu}}{1-\bar{\mu}}\frac{\gamma - g}{g}$$

Since $\bar{\mu} < \frac{g}{\gamma}$, the above expression is always less than 1.

Short-lived player 1 Finally, short-lived player 1's gain from playing D is

$$g - \gamma(\bar{\mu} + (1-\bar{\mu})\eta)$$
$$= \gamma\left\{\frac{g}{\gamma} - \bar{\mu}(1+\gamma-g)\right\}$$

This is nonnegative if

$$\frac{g}{\gamma} \geq \bar{\mu}(1+\gamma-g)$$
$$\iff 1 \geq \bar{\mu}\frac{\gamma}{g}(1+\gamma-g),$$

which is the assumption of the theorem.

□

While we consider the two-period model to illustrate the idea in a simple setting, as the time-horizon becomes longer, even a smaller probability of commitment type suffices to sustain an equilibrium in which defections are punished. The model under applied setting is discussed in Isogai and Shen (2017).

Finally, we show that the equilibrium strategy is rationalizable.

Let

$$\Delta_c = \mathcal{C}_0 \times \left\{ \mu = \{\mu_s\}_{s=1,2} : \sum_{\sigma_{-s}} \mu_s(\tau^c, \sigma_{-s}) = \bar{\mu} \right\}^2$$

be the set of CPSs in which the short-lived players believe in the ex-ante probability of the commitment type.

Theorem 6 Consider the two-period game with commitment type with $\Delta = \Delta_c$. The sets of Δ-rationalizable strategies are

$$\begin{aligned} \Sigma_0(\Delta) &= \{\sigma_0 : \ \sigma_0(\tau^c, h) = W \text{ and } \sigma_0(\tau^n, h_2) = T\} \\ \Sigma_1(\Delta) &= \Sigma_1 \\ \Sigma_2(\Delta) &= \Sigma_2 \end{aligned}$$

In particular, the PBE constructed above consists of the strategies that are extensive-form rationalizable.

Proof of Theorem 6 In the first round, the long-lived player's dominated strategies are deleted: $\Sigma_0(\Delta, 1) = \{\sigma_0 : \ \sigma_0(\tau^c, h) = W \text{ and } \sigma_0(\tau^n, h_2) = T\}$. The short-lived players' strategies are not deleted. In the second round, the long-lived player's strategies are not deleted. For the short-lived players, again no strategies is deleted. In particular, none of short-lived player 2's strategies are deleted because for a history h_2,

- if $\sigma_2(h_2) = D$, then consider μ_2 which puts probability 1 that $\sigma_0(\tau^n, h_2) = T$.
- if $\sigma_2(h_2) = C$, then consider μ_2 which puts probability 1 that the long-lived player is of commitment type.

Thus the deletion procedure ends in the second round and the result follows. □

The model considered here is a modification of Kreps and Wilson (1982)'s model to imperfect (public) monitoring.[10] The difference is that in the current model, we assume that the second short-lived player cannot observe the action of the first

[10]Also see Milgrom and Roberts (1982) and Kreps, Milgrom, Roberts, and Wilson (1982).

short-lived player. Thus the second short-lived player cannot condition his response on whether the first short-lived player has actually defected. Intuitively, this means that the second short-lived player cannot see the “reason” of the long-lived player’s action (e.g., if the long-lived player chose T, then the second short-lived player cannot distinguish between that the first short-lived player did not defect and that the long-lived player failed to punish the deviant short-lived player). Still, the equilibrium constructed here is the same as in what we would obtain under perfect-monitoring. What is affected by the imperfect-monitoring assumption is that the long-lived player might have an incentive to choose W after the first short-lived player choose C to induce the second short-lived player to choose C as well. This possibility is not an issue in our equilibrium because the long-lived player is indifferent between T and W. Therefore, I claim here that while the assumption exhibits a minor departure from Kreps-Wilson model, what we obtain as an equilibrium is the same.[11] The reason I put this analysis here is that we want to illustrate the possibility of strategic use of punishment as a signaling device when there is payoff uncertainty in stage games—other specification of payoff uncertainty also suffices.

3.5 Discussion on $\triangle$-rationalizability

One contribution in Battigalli and Sinishcanchi (2003) is that their Δ-rationalizability procedure implies iterated intuitive criteria (Cho and Kreps, 1987), which selects strategically stable equilibrium in signaling games. Consider the celebrated beer-quiche example: player 0 is of “weak” or “strong” type, and chooses between drinking beer (B) or eating quiche (Q). The strong type prefers B while the weak type prefers Q. After observing the choice by player 0, the player 1 chooses whether to “fight” (F) or “not fight” (NF). The player 1 prefers F against the weak type but prefer NF against the strong type.

The typical application of intuitive criteria to the beer-quiche game is that when the ex-ante probability of the strong type is sufficiently high so that if both types pool their actions, then player 1 prefer not to fight. In this case, there are two pure

[11]For the “reputation” model under imperfect monitoring, Fudenberg and Levine (1992) considers infinitely repeated game. To the best of my knowledge, there is no journal article that directly considers a reputation model of finitely repeated games with imperfect monitoring while Mailath and Samuelson (2006, Section 17.3) deals with an example similar to the current model.

strategy PBE, one in which both types choose B and one in which both choose Q. The intuitive criteria works as follows: in the latter equilibrium, B for weak type is an inferior response and thus B will never be chosen by weak type. Hence the rational belief for player 1 should be attaching probability 1 to strong type after observing B, which overturns the equilibrium since now the strong type prefers B to Q. Therefore, the intuitive criteria selects only the first equilibrium.

In our random-player model we can interpret "long-lived player with small realization n" as "weak" and "long-lived player with large realization n" as "strong" type. Punishment is interpreted as drinking beer while tolerating is interpreted as eating a quiche. Cooperate can be interpreted as fighting while defecting can be interpreted as not fighting. The difference here is that being a strong type does not directly affect the short-lived players' payoffs. The short-lived players' payoff depends on the long-lived player's response in the stage game, not by the long-lived player's type. Thus, as long as the long-lived player's action "tolerate" is rationalizable, so is short-lived players' action defect. This is why Δ-rationalizability cannot pin down the punishment result without the belief restriction. In summary, while our random-player model has similar structure as the beer-quiche model, the fact that there is no payoff-uncertainty within stage games makes a substantial difference in the result.

For our payoff-uncertainty case, the commitment type of the long-lived player can be directly interpreted as strong type. Still, there are two departures from the Cho-Kreps' beer-quiche example for which intuitive criterion has a bite. First, the payoff function for the commitment type takes an extremely lower payoff from tolerating a deviation; there is no equilibrium in which the commitment type of long-lived player chooses to tolerate short-lived player's defection. Second, we assumed that the probability of the commitment type is small (i.e., $g > \bar{\mu}\gamma$) so that we do not have a pooling equilibrium.

Appendix A
Construction of Harsanyi Type Space

In this Appendix, we construct the type space with which each specialized firm's belief on the number of cartels is identical and is given by $d(n),\ n = 1, 2, \ldots$ While impossibility of such symmetric belief in random-player games with unbounded support is proved in Milchtaich (2004) for proper priors, we can construct such model by using an improper prior.

Let $\mathbb{Z}$ be the set of integers and $\mathbb{N} = \{1, 2, 3, \ldots\}$ be the set of natural numbers. Our interpretation of $\mathbb{Z}$ is the set of (the identities of) potential players. The state space is $\Omega = \mathbb{Z} \times \mathbb{N}$ and the typical element (x, n) means that the set of active players is $\{x, x+1, x+2, \ldots, x+n-1\}$; in particular, the number of active players is n at state (x, n). Now assume that n is determined following the distribution $q(n) = \frac{d(n)}{n}$ and, independent of the realization of n, x is chosen according to the improper uniform distribution over $\mathbb{Z}$. Now, conditional on the player z being active, the interim belief about the number n of cartels is

$$\#\{x : \ x \leq z \leq x + n - 1\} \cdot \left(\frac{d(n)}{n} \right) = d(n).$$

Appendix B
Deterrence in The Case of Observable Actions

In this Appendix, I show that there exists an equilibrium in which cooperation is possible if we relax the unobservability assumption of the past plays. While this case is not our main focus of the paper, which takes the cartels formed by multiproduct firms as a leading example, I provide the result for the completeness of the argument and possible usefulness for other examples. The set of short-lived player s's histories is

$$\mathcal{H}_s := \{(a^1, b^1, a^2, b^2, \ldots, a^{t-1}, b^{t-1}) : \ a^\tau \in \{C, D\}, \ b^\tau \in \{T, W\}\}$$

Since the short-lived players know that whether a defection is punished or not, we can support an outcome in which punishment is credible. In particular, by stipulating that the players follow the no-punishment strategy profile after a failure to punish is observed, the long-lived player has an incentive to punish a deviator.

Theorem 7 Let $\bar{I}$ be defined as above. If $g \leq \gamma\xi^{\bar{I}}$ holds, then there is a PBE in which the punishment is used to deter deviations. If $g > \gamma\xi^{\bar{I}}$, then there is no PBE in which punishment occurs.

Proof of Proposition 7 For history

$$h = (n, a^1, b^1, \ldots, a^{t-1}) \text{ or } h = (a^1, b^1, \ldots, a^{t-1}, b^{t-1}),$$

we say that h exhibits a failure in punishment if there exists a τ such that $a^\tau = D$

and $b^\tau = W$. Otherwise we say that h exhibits no failure in punishment. Consider the following strategy profile:

Long-lived player:

- If h exhibits no failure in punishment, follow $\sigma_0^{\bar{I}}$.
- If h exhibits a failure in punishment, always choose T.

Short-lived player:

- If h exhibits no failure in punishment, then choose C.
- If h exhibits a failure in punishment, then choose D.

We show that this strategy profile constitutes a PBE.

Short-lived players:

When h exhibits a failure in punishment, the optimality is obvious. When $h = (n, \ldots, D)$ exhibits no failure in punishment, given the long-lived player's strategy, choosing D triggers the long-lived player's punishment if and only if $S_t(h^t) + I \leq n$ (note that $S_t(h^t)$ can be computed by short-lived players' information). The probability of that event is ξ^I. Thus the best response for the short-lived players is C.

Long-lived player: When h exhibits a failure in punishment, the optimality is obvious. When $h = (n, \ldots, D)$ exhibits no failure in punishment, then

1. if $S_t(h^t) + \bar{I} \leq n$, then the payoff from choosing W is $\bar{I}(1 + l) - \alpha > 0$.
2. otherwise, by $(\bar{I} - 1)(1 + l) - \alpha < 0$, choosing T is optimal.

□

The PBE constructed above survives extensive-form rationalizability.

Theorem 8 Consider the game with observable actions and $\Delta = \bar{\Delta}$. Let $\bar{I}$ be as given above. Then, if $g \leq \gamma\xi^{\bar{I}}$, then the sets of Δ-rationalizable strategies for the players are

$$\Sigma_0(\Delta) = \Sigma_0 \setminus \tilde{\Sigma_0}$$

$$\Sigma_s(\Delta) = \Sigma_s \text{ for } s = 1, 2, \ldots.$$

If $g \leq \gamma \xi^{\bar{I}}$, then the sets of Δ-rationalizable strategies for the players are

$$\Sigma_0(\Delta) = \{\mu_0 : \ \mu_0(h) = T \text{ for any } h\}$$
$$\Sigma_s(\Delta) = \{\mu_s : \ \mu_s(h) = D \text{ for any } h\} \text{ for } s = 1, 2, \ldots.$$

The proof is almost the same as Theorem 4 and is omitted.

Appendix C
Obtaining Credible Punishment under Rationalizability

Under the equilibrium assumption, the short-lived players know which actions the other short-lived players choose and thus knows that if only one short-lived player defects, then the long-lived player does not have an incentive to punish him. Generally speaking, equilibrium assumption is reasonable in many game-theoretic applications and captures the regularity of events, which we economists are mostly interested in. However, I regard this assumption might not be necessarily plausible in the current application. That the game is in equilibrium means that each player, especially each short-lived player, perfectly knows which actions the other players will take, even though they do not know even each other's existence. The analysis above tells us that if unilateral deviation by a short-lived player is possible, then he can get away from punishment.

I believe that the appropriate solution concept for analyzing the current setting is rationalizability rather than equilibrium. Here I give a set of belief restrictions which pins down the rationalizable strategies under which punishment is credibly used by the long-lived player. The model considered is the random-player game in which only occurrences of past punishment is observable as discussed in Section 3.1-3.3.

Long-lived player: If there is a short-lived player playing D, then the long-lived player believes that the following short-lived players would also choose D with positive probabilities unless there is a punishment. Moreover, the long-lived player believes that a punishment will increase the probability of short-lived players cooperating for the next $\hat{I}$ players. Formally, let $\hat{I}$ be some positive integer satisfying $\hat{I}(1+l) > \alpha$;

$$\tilde{\Sigma}_s = \left\{\sigma_s : \ \sigma_s(b^1, \ldots, b^{s-1}) = C \text{ if and only if } b^t = W \text{ for some } t \geq s - \hat{I}\right\}$$

be the set of short-lived player's strategies which plays C if and only if a punishment occurred in the past $\hat{I}$ periods. Define the belief restriction

$$\tilde{\Delta}_0 = \left\{\mu_0 : \ \mu_0\left(\prod_{s\geq 1} \tilde{\Sigma}_s \middle| h_0\right) \geq \pi \text{ for all } h_0 \in \mathcal{H}_0\right\}.$$

Short-lived player: Once a long-lived player chooses W, then the short-lived players will believe that the long-lived player would follow σ_0^I with probability 1:

$$\tilde{\Delta}_s = \left\{\mu_s : \ \sum_{\sigma_{-s}} \mu_s(\sigma_0^{\hat{I}}, \sigma_{-s}|h_s) = 1 \text{ if } \exists\tau, \ b^\tau = W \text{ and } h_s = (b^1, \ldots, b^s)\right\}$$

Theorem 9 Consider the baseline random-player game with $\Delta = \bar{\Delta} \cap (\tilde{\Delta}_0 \times \prod_s \tilde{\Delta}_s)$. Assume $\pi\hat{I}(1+l) > \alpha$ and $g < \gamma\xi^{\hat{I}}$. The sets of Δ-rationalizable strategies are[1]

$$\Sigma_0 = \{\sigma_0 : \ \sigma_0(n, a^1, b^1, \ldots, a^t) = W \text{ iff } a^\tau = C, \ 1 \leq \tau \leq t-1; \ a^t = D; \ t + \hat{I} \leq n\}$$
$$\Sigma_s = \{\sigma_s : \ \sigma_s(h) = C \text{ for } h = (b^1, \ldots, b^t) \text{ with } b^\tau = W \text{ for some } \tau\}$$

In particular, under Δ-rationalizable strategies, if a short-lived player s chooses D for the first time and $s + \hat{I} \leq n$, then the long-lived player chooses W, after which (C, T) follows in the rest of the game.

<u>**Proof of Theorem 9**</u> The deletion procedure ends in the third round

First round For the long-lived player, consider the history $h = (n, a^1, b^1, \ldots, a^t)$ such that $a^\tau = C$ and $b^\tau = T$ for $\tau = 1, \ldots, t-1$, $a^t = D$, and $t + \bar{I} \leq n$. By the belief restriction, choosing T is dominated.

[1] "iff" means "if and only if."

For the short-lived players, consider the history $h = (b^1, \dots, b^{t-1})$ such that there exists a $1 \leq \tau \leq t-1$ such that $b^\tau = W$. By the belief restriction, the action D is dominated.

Second round Given $\Sigma_s(\Delta, 1)$, choosing W in histories other than above is dominated for the long-lived player.

□

While the proof is almost trivial, the belief restriction suggests a possible realistic story. First, the long-lived player might infer an occurrence of deviation as a symptom for other short-lived players to deviate since the short-lived players are symmetric and, thus, once one short-lived player had an incentive to deviate, it is natural to believe that others will also deviate unless there is no further punishment. The long-lived player's belief restriction also requires that the long-lived player believes that an occurrence of punishment successfully affects the short-lived players' actions.

The second point, which might be more realistic is that an occurrence of punishment without a doubt signals the short-lived players that there was certainly a punishment, while non-existence of punishment is not distinguishable between the non-existence of deviation and failures to punish. Once a punishment occurs, the short-lived players come to know the intention of the long-lived player to punish deviators and thus the punishment becomes credible.

Appendix D
Tables for Notations in Chapter 3

Table D.1. Notations in Section 3.3

$d(n)$	(geometric) probability distribution over the number of short-lived players
u^1, u^2	stage-game payoff functions
$\hat{U}$	extensive-form payoff function
U	strategic-form payoff function
$\mathcal{H}_i$	set of player i's non-terminal histories
$\mathcal{Z}$	set of terminal histories (outcomes)
Σ_i	set of player i's strategies
ζ	outcome function which maps strategy profiles to outcomes
$\mathcal{C}_i$	set of conditional probability systems for player i
$\bar{\Delta}$	belief restrictions for players at which all short-lived players' beliefs are consistent with ex-ante probability $d(n)$
$\Sigma_i(\Delta, n)$	set of player i's strategies which survive n-th round of Δ-rationalizability procedure
$\Sigma_i(\Delta)$	set of player i's Δ-rationalizable strategies
$\bar{I}$	the smallest number of remaining short-lived players to justify the use of punishment; i.e., the integer that satisfies $(\bar{I}+1)(1+l) \geq \alpha > \bar{I}(1+l)$.
$\tilde{\Sigma_0}$	the set of long-lived player's strategies which stipulates punishment when there are not enough short-lived players; i.e., $\tilde{\Sigma_0} = \{\sigma_0 : \ \sigma_0(h_0) = W \text{ if } h_0 \in \mathcal{H}^t(n) \text{ and } t + \bar{I} > n\}$

Table D.2. Notations in Section 3.4

$\bar{\mu}$	probability of the long-lived player being the commitment type
$\mathcal{T}$	payoff-type space $\{\tau^n, \tau^c\}$ of the long-lived player
$\hat{\mathcal{H}}_i$	set of non-terminal histories for player i in the game with payoff uncertainty
$\hat{\mathcal{Z}}$	set of terminal histories in the game with payoff uncertainty
$\hat{U}_i$	extensive-form payoff function for player i in the game with payoff uncertainty
$\Sigma_i(\Delta, n)$	set of player i's strategies which survive n-th round of Δ-rationalizability procedure
$\Sigma_i(\Delta)$	set of player i's Δ-rationalizable strategies

Bibliography

[1] Battigalli, Pierpaolo (1997), “On Rationalizability in Extensive Games,” *Journal of Economic Theory*, Volume 74, Issue 1, Pages 40-61.

[2] Battigalli, Pierpaolo (2001), “A Note on Rationalizability and Reputation with Two Long-Run Players,” *RISEC (International Review of Economics and Business)*. nr. 48, pp. 383-393.

[3] Battigalli, Pierpaolo (2003), “Rationalizability in infinite, dynamic games with incomplete information,” *Research in Economics*, 57, 1?38

[4] Battigalli, Pierpaolo and Amanda Friedenberg (2012), “Forward induction reasoning revisited.” *Theoretical Economics* 7, 57-98.

[5] Battigalli, Pierpaolo and Marciano Siniscalchi (2002), “Strong Belief and Forward Induction Reasoning,” *Journal of Economic Theory*, Volume 106, Issue 2, Pages 356-391.

[6] Battigalli, Pierpaolo and Marciano Siniscalchi (2003), “Rationalization and incomplete information.” *Advances in Theoretical Economics*, 3(1).

[7] Benoît, Jean-Pierre, and Vijay Krishna (1993), “Renegotiation in Finitely Repeated Games,” *Econometrica*, Vol. 61, No. 2, pp. 303-323.

[8] Ben-Porath, Elchanan and Dekel, Eddie (1992), “Signaling Future Actions and the Potential for Sacrifice”, *Journal of Economic Theory*, 57(1), 36-51.

[9] Bernheim, B. Douglas and Erik Madsen (2017) “Price Cutting and Business Stealing in Imperfect Cartels,” *The American Economic Review*, 107(2), 387-424.

[10] Bernheim, B. Douglas and Debraj Ray (1989), “Collective dynamic consistency in repeated games,” *Games and Economic Behavior*, Volume 1, Issue 4,, Pages 295-326.

[11] Chen, Jing, and Silvio Micali (2013), “The order independence of iterated dominance in extensive games,” *Theoretical Economics*, 8, 125-163.

[12] Cho, In-Koo, and David M. Kreps (1987), “Signaling games and stable equilibria.” *Quarterly Journal of Economics*, 102(2), 179-222.

[13] Farrell, Joseph, and Eric Maskin (1989), “Renegotiation in repeated games,” *Games and Economic Behavior*, 1(4), 327-360

[14] Fudenberg, Drew, and David K. Levine“Maintaining a Reputation when Strategies are Imperfectly Observed,” *Review of Economic Studies*, Vol. 59, pp. 561-579.

[15] Fudenberg, Drew and Jean Tirole (1991), “Game Theory,” *The MIT Press.*

[16] Govindan, Srihari and Robert Wilson (2009), “On Forward Induction,” *Econometrica*, Volume 77, Issue 1, Pages 1?28.

[17] Isogai, Shigeki (2016) “Credibility of Deterrence and Extensive-Form Rationalizability,” *mimeo.*

[18] Isogai, Shigeki (2017) “Credibility of Deterrence and Extensive-Form Rationalizability: Infinite-Horizon Version,” *mimeo.*

[19] Isogai, Shigeki and Chaohai Shen (2017) “Multiple Cartels and Strategic Use of The Leniency Policy,” *mimeo.*

[20] Kohlberg, Elon, and Jean-Francois Mertens (1986), “On the Strategic Stability of Equilibria,” *Econometrica*, 54(5), 1003-1037

[21] Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson (1982), “Rational Cooperation in the Finitely Repeated Prisoners’ Dilemma,” *Journal of Economic Theory*, 27, 245-252.

[22] Kreps, David M., and Robert Wilson (1982), “Reputation and imperfect information,” *Journal of Economic Theory*, Volume 27, Issue 2 Pages 253-279.

[23] Mailath, George J., and Larry Samuelson (2006), “Repeated Games and Reputations: Long-Run Relationships,” *Oxford University Press.*

[24] Milgrom, Paul, and John Roberts (1982), “Predation, Reputation, and Entry Deterrence,” *Journal of Economic Theory*, 27, 280-312.

[25] Milchtaich, Igal (2004), “Random-player games.” *Games and Economic Behavior*, 47, 353-388.

[26] Myerson, Roger B. (1997), “Game Theory: Analysis of Conflict,” *Harvard University Press.*

[27] Myerson, Roger (1998), "Population uncertainty and Poisson games." *International Journal of Game Theory*, November 1998, Volume 27, Issue 3, pp 375-392

[28] Osborne, Martin J. (1990), "Signaling, Forward Induction, and Stability in Finitely Repeated Games," *Journal of Economic Theory*, 50, 22-36.

[29] Osborne, Martin J. and Ariel Rubinstein (1994), "A Course in Game Theory," *The MIT Press*.

[30] Pearce, David G. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, Vol. 52, No. 4, pp. 1029-1050.

[31] Perea, Andres (2010), "Backward Induction versus Forward Induction Reasoning," *Games*, 1(3), 168-188.

[32] Reny, Philip J. (1992), "Backward Induction, Normal Form Perfection and Explicable Equilibria," *Econometrica*, Vol. 60, No. 3, pp. 627-649.

[33] Rênyi, A. (1955), "On A New Axiomatic Theory of Probability," *Acta Mathematica Hungarica,* 6, 285-335.

[34] Shimoji, Makoto, and Joel Watson (1998), "COnditional Dominance, Rationalizability, and Game Forms," *Journal of Economic Theory*, 83, 161-195.

[35] van Damme, Eric (1989), "Stable equilibria and forward induction," *Journal of Economic Theory*, 48(2), 476-496

Vita

Shigeki Isogai

Shigeki Isogai was born on August 22 1986 in Kyoto, Japan. He earned his B.A. degree in Economics in 2010 at Kyoto University, Japan. He earned his M.A. degree in Economics in 2012 at Kyoto University, Japan. Since then, he has continued his studies at Penn State University as a graduate student. His Ph.D. thesis has focused on applications of epistemic game theory to economic problems such as cartels.