Kyoto University,
Graduate School of Economics
Discussion Paper Series

Nash Reversion Revisited: Implications of Gain/Loss Asymmetry

Tadashi Sekiguchi and Katsutoshi Wakai

Discussion Paper No. E-24-009

Graduate School of Economics
Kyoto University
Yoshida-Hommachi, Sakyo-ku
Kyoto City, 606-8501, Japan

March, 2025

Nash Reversion Revisited: Implications of Gain/Loss Asymmetry*

Tadashi Sekiguchi† and Katsutoshi Wakai‡

March 22, 2025

Abstract

Game theory proves the existence of a stronger punishment than the Nash reversion in the repeated games. Recent empirical findings in Oligopoly, however, suggest the implementation of the Nash reversion. In a standard repeated game setting, we propose a potential answer for this empirical puzzle by using a refined version of the discounted utility that exhibits gain/loss asymmetry, where players discount gains more than losses. Our main result is as follows: among gain/loss robust subgame perfect equilibria, the Nash reversion offers the strongest punishment. The robustness is based on the assumption that players are unsure about their own level of gain/loss asymmetry and choose only the strategies that are subgame perfect for any level of gain/loss asymmetry they can perceive as possible.

Keywords: Gain/loss asymmetry, optimal penal code, repeated game, recursive utility, utility smoothing

JEL Classification: C73, D20, D90, L13

---

*This paper is a revised version of the discussion paper (E-16-006, Graduate School of Economics, Kyoto University) entitled as "Repeated Games with Recursive Utility: Cournot Duopoly under Gain/Loss Asymmetry."

†Kyoto Institute of Economic Research, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan; E-mail: sekiguchi@kier.kyoto-u.ac.jp

‡Graduate School of Economics, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan; E-mail: wakai@econ.kyoto-u.ac.jp

1 Introduction

1.1 Overview

Over the decades, theoretical work on repeated games have confirmed that a choice of credible punishment is a crucial factor for determining a level of collusion. A well-known punishment path is a sequence of the static Nash equilibrium action, and a trigger strategy profile that uses this sequence as punishment is called the Nash reversion strategy profile. As for the strongest punishment, Abreu (1988) shows in his seminal work that the worst subgame perfect equilibrium, referred to as the optimal penal code, can be used as punishment. Abreu (1986) identifies this optimal penal code for repeated Cournot oligopoly and demonstrates that it can be used to implement a collusive equilibrium path that is not sustainable under the conventional Nash reversion strategy profile.

With Abreu' (1986, 1988) results, we expect that firms have a strong incentive to use the optimal penal code in an actual oligopolistic situation. However, in a recent study of the well-known international Oligopoly formed in the Vitamin markets, Igami and Sugaya (2022) confirm that the data, as well as the comments by cartel participants on legal documents, supports the implementation of the Nash reversion strategy profile. Their finding raises a puzzle: why do firms choose the Nash reversion strategy profile over alternative equilibrium strategy profiles that seem to support a better equilibrium path?

We conjecture that this puzzle is a consequence of the discounted utility, which is a standard assumption in almost all repeated games without complete patience. Our claim is that we can propose a potential answer for this puzzle if we use a model that extends the discounted utility consistently to some of its empirical anomalies. The anomaly we consider to be a cause of this puzzle is gain/loss asymmetry, which is a phenomenon whereby a decision maker tends to discount gains more than he discounts losses. This anomaly was first documented by Thaler (1981) and repeatedly confirmed in many subsequent experiments.[1]

Among a few representations that are capable to induces gain/loss asymmetry, we choose a version of the preference representation suggested by Wakai (2008), which identifies a notion of intertemporal utility smoothing as a source of gain/loss asymme-

[1] See Frederick, Loewenstein, and O'Donoghue (2002) for a survey of experimental studies on gain/loss asymmetry.

try. The advantage of Wakai' (2008) representation is that it belongs to a class of the recursive utility suggested by Koopmans (1960) that describes history-independent, stationary, and dynamically consistent preferences. In an oligopolistic situation, firms are expected to be sophisticated enough to behave rationally so that the assumption of recursivity is central. Moreover, we want to resolve the puzzle with the minimal departure from the discounted utility.

Formally, at each time t, some of the players evaluate a payoff sequence $U = (u_0, u_1, ...)$ based on the following function:

$$
\begin{aligned}
V_t(U) &= V_t((u_t, u_{t+1}, ...)) \\
&\equiv \min_{\delta \in [\underline{\delta}, \overline{\delta}]} [(1-\delta)u_t + \delta V_{t+1}(U)] \\
&= u_t + \underline{\delta} \max \{V_{t+1}(U) - u_t, 0\} + \overline{\delta} \min \{V_{t+1}(U) - u_t, 0\}.
\end{aligned}
\tag{1}
$$

where $\underline{\delta}$ and $\overline{\delta}$ are parameters satisfying $0 < \underline{\delta} \leq \overline{\delta} < 1$. Evaluating function (1) is one-parameter richer than the discounted utility, which is a special case of (1) with $\underline{\delta} = \overline{\delta}$. The value of (1) also becomes a weighted average of a payoff sequence at the effective selection of discount factors, where an exact form of the representation will be explained in the text.

Evaluating function (1) exhibits a key feature called *recursive gain/loss asymmetry*: the difference between future value $V_{t+1}(U)$ and current payoff u_t defines a gain or loss, and gains and losses are discounted by $\underline{\delta}$ and $\overline{\delta}$, respectively. Thus, gains are discounted more than losses. We call $\underline{\delta}$ and $\overline{\delta}$ the gain discount factor and the loss discount factor, respectively, whereas the distance between $\underline{\delta}$ and $\overline{\delta}$ represents a degree of gain/loss asymmetry. Moreover, it has been shown that the player j is said to be *more time-variability averse* than the player i if $[\underline{\delta}_j, \overline{\delta_j}] \supseteq [\underline{\delta}_i, \delta_i]$. A more time-variability averse player prefers less volatile utility (equivalently, payoff) sequences.

Besides the experimental motivation mentioned above, gain/loss asymmetry seems to be consistent with observed behavior of firms. To sustain the oligopoly, firms must be sensitive enough to avoid future punishment, which suggests the usage of a high discount factor when they face a future loss. On the other hand, the corporate finance literature has documented a short-term oriented nature of managers' objectives. With a notion of loss aversion in mind, this tendency must be more evident if current payoff is regarded as a loss relative to the value receiving in the future, which suggests the usage of a low discount factor when they face a future gain.

Given the above intuition, we first investigate the properties of the optimal penal code for the recursive utility because it is the largest class of preference relations that inherit the key features assumed in the discounted utility. This also helps us to isolate properties unique to (1). We confirm that the implications derived by Abreu (1988) for general stage games and by Abreu (1986) for a particular application to repeated Cournot oligopoly extend to the recursive utility. For the latter implication, the optimal penal code has a structure similar to a stick-and-carrot structure shown in Abreu (1986): a player deviating from a target outcome path will be penalized severely until it accepts the penalty, and this deviator will receive a better payoff as compensation in a subsequent period once the deviator accepts the penalty.

With this in mind, we now investigate the limiting behavior of the optimal penal code, where the gain discount factor $\underline{\delta}$ converges to zero while a loss discount factor $\overline{\delta}$ is given and fixed. This investigation is possible only for evaluating function (1) because in the discounted utility model, there is only a single discount factor. We then identify the property, referred to as a *reverse* Folk theorem, which is new and unique to evaluating function (1): Consider the set of pure strategy subgame perfect equilibria in the repeated game that satisfy a certain regularity condition. Assume that one of the players preferences are represented by (1). As the gain discount factor $\underline{\delta}$ goes to zero for this player, that is, as the player becomes extremely averse to a current period loss, the penal code used in the Nash reversion strategy profile *asymptotically* becomes an optimal penal code. The regularity condition we impose is called a *coherent action space*, which generalizes the properties shown in the symmetric subgame perfect equilibria. Therefore, our result covers some of asymmetric subgame perfect equilibria.

Evaluating function (1) is a form of recursive utility, where the optimal penal code exhibits the stick-and-carrot structure. Because the current utility is weakly less than the utility of future payoffs, the gain discount factor $\underline{\delta}$ is used to evaluate the future utility. As a decrease in the gain discount factor $\underline{\delta}$ causes the decreased anticipation of future gain, the player no longer accepts a sever current penalty at a given level of future compensation. Therefore, the optimal penalty becomes weaker. This suggests that the gain discount factor $\underline{\delta}$ controls the *enforceability* of the optimal penal code. Intuitively, our result states that when the enforceability of the optimal penal code vanishes, the current action in the optimal penal code must converge to the Nash equilibrium action.

In terms of the role of the loss discount factor $\overline{\delta}$, we investigate the limiting behavior of equilibria where a loss discount factor $\overline{\delta}$ converges to one while a gain discount factor $\underline{\delta}$ is given and fixed. We show that as the loss discount factor $\overline{\delta}$ approaches to one, the player increasingly fears the penalty so that the player avoids deviating from the collusive path. Thus, the loss discount factor $\overline{\delta}$ determines the level of *deterrence* of a given penal code. Moreover, this result is independent of the gain discount factor $\underline{\delta}$. Therefore, an increase in the loss discount factor $\overline{\delta}$ opens a possibility of collusion even if the player is nearly myopic when facing a current period loss, that is, when the gain discount factor $\underline{\delta}$ is nearly zero.

We now state our main result: the penal code used in the Nash reversion strategy profile is a unique optimal penal code when we restrict an attention to the collection of the subgame perfect equilibria that are robust to the gain/loss asymmetry in discount factors. Here, a robust equilibrium is defined as a strategy profile that is subgame perfect for any combination of players who are more time-variability averse than the players with given thresholds of gain and loss discount factors. This assumption captures a situation where players are unsure about a level of gain/loss asymmetry except the threshold level and avoid selecting a strategy profile that is an equilibrium only under a particular combination of discount factors. In this sense, we effectively introduce a selection criteria for subgame perfect equilibria of repeated games.

To study the robust equilibria discussed above, it is essentially identical to consider the strategy that remains as an equilibrium when the players become more time-variability averse. Thus, a decrease in the gain discount factor $\underline{\delta}$ forces the optimal penal code to converges to the strategy that generates a sequence of the static Nash equilibrium action, while an increase in the loss discount factor $\overline{\delta}$ makes the player more foresighted so that collusive payoffs can be attained. We confirm that when the players use the Nash reversion strategy profile, a constant sequence that is an equilibrium under the threshold level of gain/loss asymmetry remains to be an equilibrium for a more time-variability averse player. This result contrast to the findings in a standard discounted utility: the set of equilibria is simply getting smaller if discount factors decreases. Thus, only the strategy that generates a sequence of the static Nash equilibrium action remains at the limit, and the players no longer obtain any collusive payoffs.

1.2 Outline and Related Literature

The paper proceeds as follows. Section 2 defines the setting of the game and examines the characteristics of evaluating function (1). Assuming a general form of recursive utility, Section 3 characterizes the optimal penal code and the associated reversion strategy profile. Section 4 defines the coherent action space and derives the reverse Folk theorem. We also examine the relation between evaluating function (1) and the standard Folk theorem. Section 5 shows our main result that among the gain/loss robust subgame perfect equilibria, the Nash reversion strategy profile is based on the strongest punishment. Section 6 concludes the paper with Section 7 acknowledging the support. Most of proofs are presented in the appendices.

We now review the related literature. As for the models of discounting that exhibit gain/loss asymmetry, evaluating function (1) is closely related to the models suggested by Loewenstein and Prelec (1992) and Shalev (1997), both of them are motivated by loss aversion (Kahneman and Tversky (1979), Tversky and Kahneman (1991)). Loewenstein and Prelec (1992) define a gain or a loss based on a variation in a utility sequence, whereas Shalev (1997) defines a gain or a loss based on the difference in utility between adjacent periods. However, the aforementioned models describe static choices, and it does not belongs to recursive preference.

In terms of consistency with empirical findings, Igami and Sugaya (2022) estimate the repeated game implications based on a more realistic setup that includes a demand shifter and competitive fringe suppliers. In their model, the firms engaged in the cartel first expected no break up on the equilibrium path, but the cartel actually broke down on the course because unforeseen negative news about competitive fringe suppliers arrived. Our model is regarded as an approximation of the cartel phase as well as the broke down phase, where in the latter phase firms produced the Nash equilibrium quantity.

Some papers study repeated games without the assumption of discounted utility. Chade, Prokopovych, and Smith (2008) assume that the players have β–δ preferences, and characterize the Strotz-Pollak equilibria by Peleg and Yaari (1973). Obara and Park (2017) extend the analysis to general discounting functions including the β–δ ones as special cases.[2] We explore the implications of general recursive utility.

In that sense, their companion paper (Obara and Park (2014)) is more closely related to ours because it considers a class of time preferences including recursive utilities as special cases. The authors' results on the structure of Strotz-Pollak equi-

[2]The β–δ preferences are a class of the discount functions with present bias. Obara and Park (2017) consider both the discounting functions with present bias and the functions with future bias.

libria are similar to our results under recursive utility.[3] The difference is that while Obara and Park (2014) examine how the general methodology for the discounted utility model extends to general time preferences possibly without recursive structure (under the Strotz-Pollak equilibrium concept), we focus on the case of recursive gain/loss asymmetry and provide its implications on the determination of the optimal penal code in detail. Their paper also provides a proof for the existence of the equilibria in their context, where if it is restricted to the recursive utility, it can be regarded as the proof of Proposition 1.

Kochov and Song (2023) employ the recursive utility criterion by Uzawa (1968) and Epstein (1983) as well as the recursive utility criterion by Epstein and Zin (1989). In the former model, a player's discount factor for the next period depends on the current period payoff, whereas in the latter model, the discount factor is fixed for the deterministic sequences. They find situations for the players to manage risk and trade intertemporally. In our model, we focus on pure strategy subgame perfect equilibria to derive clear implications for the optimality of the Nash reversion strategy profile via the role of gain/loss asymmetry.

2 Model

2.1 Setting

The stage game, denoted by $G = \left(I, \{A_i\}_{i=1}^{I}, \{u_i\}_{i=1}^{I}\right)$, is an I-player simultaneous move game, where player i's action space A_i is a compact metric (therefore, topological) space with multiple elements and player i's payoff (possibly atemporal utility) function $u_i : \prod_{i=1}^{I} A_i \to \mathbb{R}$ is continuous.[4] By abuse of notation, I indicates both the number of players and the set of the players. We define A by $A \equiv \prod_{i=1}^{I} A_i$ and use $q^{(t)} \equiv (q_1^{(t)}, ..., q_I^{(t)}) \in A$ to denote a vector of actions taken at time t by all players, where time t varies over $\mathbb{N} = \{0, 1, 2, ...\}$. We refer to $q^{(t)}$ as an action profile or a time-t action profile if we want to emphasize the time period. We also use $q_{-i}^{(t)}$ to indicate the time-t action profile by all players except player i.

[3] The Strotz-Pollak equilibrium concept reduces to the standard equilibrium concept under recursive utility.

[4] As usual, the distance is not zero between different elements in A.

We consider the supergame G^{∞} obtained by repeating game G infinitely often. For each t with $t > 0$, let H^{t-1} be defined by $H^{t-1} \equiv A^{t}$, each element of which, $h^{t-1} \equiv (q^{(0)}, ..., q^{(t-1)})$, is a series of realized actions at all periods before period t. We assume that $H^{-1} \equiv A^{0}$, which consists of a single element. For each player i, we focus on a pure strategy s_i, that is, a sequence of functions $s_i \equiv \{s_{i,t}\}_{t=0}^{\infty}$, where $s_{i,t} : H^{t-1} \to A_i$ for each t. A strategy profile s is defined by $s \equiv (s_1, ..., s_I)$, and let S be the collection of strategy profiles. We define the time-t action profile $q^{(t)}(s)$ by $q^{(t)}(s) \equiv (q_1^{(t)}(s), ..., q_I^{(t)}(s))$, where $q_i^{(t)}(s)$ is the time-t action taken by player i when the players follow s. A path Q is a sequence of action profiles denoted by $Q \equiv (q^{(0)}, q^{(1)}, ...)$. In particular, $Q(s)$ is the path of the strategy profile s, that is, $Q(s) \equiv (q^{(0)}(s), q^{(1)}(s), ...)$. Moreover, for any path $Q \in (A)^{\infty}$, let $U_i(Q)$ be the sequence of player i's payoffs $(u_i(q_1^{(0)}, ..., q_I^{(0)}), u_i(q_1^{(1)}, ..., q_I^{(1)}), ...) \in [u_i(A)]^{\infty}$, where u_i is continuous on A and $u_i(A)$ is the image of u_i.

We adopt a continuous and strictly monotone utility function $V_{i,t} : [u_i(A)]^{\infty} \to \mathbb{R}$, where we adopt the product topology on A^{∞} as well as $[u_i(A)]^{\infty}$. Because A is compact and u_i is continuous on A, it follows from Tychonoff's theorem that A^{∞} and $[u_i(A)]^{\infty}$ are compact. Moreover, the continuity of $V_{i,t}$ implies that the image of $V_{i,t}$ is compact. We then say that player i's preferences follow the recursive utility if $V_{i,t}$ satisfies

$$V_{i,t}(U_i(Q)) = W_i(u_i(q_1^{(t)}, ..., q_I^{(t)}), V_{i,t+1}(U_i(Q))), \tag{2}$$

where the aggregator function $W_i : conv\,[u_i(A)] \times \mathbb{R} \to \mathbb{R}$ is continuous and strictly increasing in both arguments. The recursive utility represents all classes of continuous and dynamically consistent intertemporal preferences, which are history-independent, stationary, and monotone in payoffs.[5] Furthermore, let $\overline{Q}(q)$ be a constant sequence of $q \in A$. We then define the *implied utility of constant payoff* $u_i(q)$ by

$$u_i^*(q) \equiv V_{i,0}(U_i(\overline{Q}(q))),$$

which is continuous on A.[6] We interpret $u_i^*(q)$ as the atemporal utility of action q expressed in a comparable scale to the utility of payoff sequences. We will frequently utilize this function in the following sections.

[5] Given history independence, $V_{i,t}$ depends solely on $(u_i(q^{(t)}), u_i(q^{(t+1)}), ...)$.

[6] An open set in A is identified with an open set in $C \subseteq A^{\infty}$ under the subspace topology, where C is the collection of all constant paths.

2.2 Recursive Preferences Exhibiting Gain/Loss Asymmetry

To show the robustness of the Nash reversion strategy profile, we later assume that some or all of the players' utility functions satisfy the following class of the recursive utility: At each time t, player i evaluates his payoff sequence $U_i(Q)$ based on a version of the model of utility smoothing as developed by Wakai (2008):

$$V_{i,t}(U_i(Q)) \equiv \min_{\{\delta_{t+\tau}\}_{\tau=1}^{\infty} \in [\underline{\delta}_i, \overline{\delta}_i]^{\infty}} \left\{ \sum_{\tau=0}^{\infty} (1 - \delta_{t+\tau+1}) \left(\prod_{\tau'=1}^{\tau} \delta_{t+\tau'} \right) u_i(q^{(t+\tau)}) \right\}, \qquad (3)$$

where $\underline{\delta}_i$ and $\overline{\delta}_i$ are parameters satisfying $0 < \underline{\delta}_i \leq \overline{\delta}_i < 1$. Thus, representation (3) leads to a weighted average of a payoff sequence, where the sequence of the weights applied depends on the nature of the payoff sequence.[7]

Representation (3) is a class of the recursive utility (2) under the aggregator function satisfying the following relation

$$W_i(u_i(q^{(t)}), V_{i,t+1}(U_i(Q))) = \min_{\delta \in [\underline{\delta}_i, \overline{\delta}_i]} [(1 - \delta) u_i(q^{(t)}) + \delta V_{i,t+1}(U_i(Q))]. \qquad (4)$$

Following the explanation in the introduction, we will call $\underline{\delta}_i$ and $\overline{\delta}_i$ the gain discount factor and the loss discount factor, respectively, whereas the distance between $\overline{\delta}_i$ and $\underline{\delta}_i$ represents the degree of gain/loss asymmetry. Moreover, representations (3) and (4) satisfy the following two properties:

(i) For any $q \in A, u_i^*(q) = u_i(q)$.

(ii) For any $q \in A, u_i^*(q) = W_i(u_i(q), u_i^*(q))$.

Under these conditions, the utility of future payment becomes comparable to the (implied) utility of the current payment.

The gain/loss asymmetry introduced in (3) expresses the desire to lower the volatility involved in a payoff sequence. To see this further, let U denote a sequence of payoffs, and let $\overline{u}$ denote a sequence of a constant payoff u. Then, by following Wakai (2008), we say that a player j is *more time-variability averse* than a player i if for any $\overline{u}$ and any U,

$V_{i,t}(\overline{u}) \geq V_{i,t}(U)$ implies $V_{j,t}(\overline{u}) \geq V_{j,t}(U)$,

and the latter is strict if the former is strict.

[7]It is easy to see that $\sum_{\tau=0}^{\infty} (1 - \delta_{t+\tau+1}) \left(\prod_{\tau'=1}^{\tau} \delta_{t+\tau'} \right) = 1$ for any t.

Two players agree on the ranking of a constant payoff sequence, whereas any payoff sequence disliked by the player i is disliked by the player j. Moreover, Wakai (2008) shows that this relation is translated into the set inclusion, that is, the player j is more time-variability averse than the player i if and only if

$$[\underline{\delta}_j, \overline{\delta}_j] \supseteq [\underline{\delta}_i, \overline{\delta}_i], \tag{5}$$

where $[\underline{\delta}_i, \overline{\delta}_i]$ and $[\underline{\delta}_j, \overline{\delta}_j]$ represent sets of discount factors for $V_{i,t}$ and $V_{j,t}$, respectively.

To show the robustness of the Nash reversion strategy profile, we later examine the situation where a player is strongly inclined to shift the future payoffs to the current period when the utility of the future payoffs is higher than the utility of the current payoff. For example, if a firm faces a steady stream of profits but encounters a shortfall in the current profit, to satisfy the stockholders request for a steady stream of dividend payment, it may well have a strong incentive to adjust production plan to smoothen the stream of profits. Then, the firm ends up with behaving nearly myopically when the current profit is lower than the future profit. This situation corresponds to the case where the firm has a significantly small gain discount factor $\underline{\delta}_i$. In Section 4, we will investigate a case where such a tendency progressively become stronger by letting the gain discount factor $\underline{\delta}_i$ approach to zero.

3 Background Results of Recursive Utility

In this section, we identify and characterize the optimal penal code under the recursive utility. This analysis highlights the consequences of the key assumptions inherited from the discounted utility and provides background results for our main investigation.

Let S^* be the set of all pure strategy subgame perfect equilibria of G^∞. We assume that S^* is nonempty.[8] Moreover, for each i, we define $\underline{v}_i$ and $\overline{v}_i$ by

$$\underline{v}_i \equiv \inf \left\{ V_{i,0}(U_i(Q(s))) \,|\, s \in S^* \right\}$$

and

$$\overline{v}_i \equiv \sup \left\{ V_{i,0}(U_i(Q(s))) \,|\, s \in S^* \right\},$$

[8] A sufficient condition applied to any W_i is that the stage game G has a pure strategy Nash equilibrium.

where $V_{i,0}$ follows (2). Appendix A shows the existence of subgame perfect equilibria $\underline{s}^i$ and $\overline{s}^i$ in S^* under which player i's utility is $\underline{v}_i$ and $\overline{v}_i$, respectively. This result is an extension of Abreu's (1988) Proposition 2 to the situation of recursive utility.

Next, for an $(I+1)$-tuple of paths $(Q, Q^1, \ldots, Q^I)$, we define a *reversion strategy profile* $s(Q, Q^1, \ldots, Q^I)$ as follows: (i) Q is the initial ongoing path, and players play it until some player deviates unilaterally from it, and (ii) if player j unilaterally deviates from the current ongoing path, Q^j becomes the next ongoing path, and they play it until some player deviates unilaterally from it.[9] If all the players follow this reversion strategy profile without a deviation, the path becomes Q regardless of whether it is an equilibrium. In Appendix A, we show how to construct $\underline{s}^i$ and $\overline{s}^i$ based on the reversion strategy profiles.

A key result of Abreu (1988) is that for the discounted utility model, any subgame perfect equilibrium path Q is implemented as a subgame perfect equilibrium by the reversion strategy profile $s(Q, Q(\underline{s}^1), \ldots, Q(\underline{s}^I))$. Under the equilibrium, the ongoing path after any player's unilateral deviation is the player's worst equilibrium path. The vector $(\underline{s}^1, ..., \underline{s}^I)$ is called an *optimal penal code*. Given this result, for an expositional purpose, we call $\underline{s}^i$ the *optimal penalty* for the player i. We also refer to a collection of subgame perfect equilibria $(s^1, ..., s^I)$ as a *penal code* if there exists some path Q such that $s(Q, Q(s^1), \ldots, Q(s^I))$ becomes the subgame perfect equilibrium in S^*. With this definition, a penal code that implements all of the subgame perfect equilibrium paths via the reversion strategy profile is an optimal penal code.

The above result simplifies the analysis of subgame perfect equilibria because we can restrict our attention to the paths that are supported by the optimal penal code. The next proposition shows that the same simplification holds for the recursive utility (see Appendix A for the proof).

Proposition 1: *Suppose that players evaluate payoff sequences by* (2)*. Then, for any subgame perfect equilibrium* s^* *in* S^**, the equilibrium path* $Q(s^*)$ *can be generated as the path of the reversion strategy profile* $s(Q(s^*), Q(\underline{s}^1), ..., Q(\underline{s}^I))$*, where* $s(Q(s^*), Q(\underline{s}^1), ..., Q(\underline{s}^I))$ *is a subgame perfect equilibrium in* S^**.*

Proposition 1 shows that the nonlinearity introduced on the aggregator function W_i does not alter the topological nature of the set of pure strategy subgame perfect

[9] This reversion strategy profile $s(Q, Q^1, ..., Q^I)$ corresponds to a *simple strategy profile* in Abreu (1988).

equilibria or the effectiveness of the reversion strategy profiles. Therefore, monotonicity, continuity, and recursivity are key properties that derive these results. Note that W_i and u_i need not be identical to W_j and u'_j if $i \neq j$, where u'_j is the function obtained from u_j by permuting ith and jth element.

Having confirmed the effectiveness of the reversion strategy profile, we next characterize the basic properties of the optimal penal code (see Appendix B for the proof). For this analysis, let $BR_i(q^{(t)}_{-i})$ denote any one of the player i's best responses in the stage game G when other players take action profile of $q^{(t)}_{-i}$, where the existence of $BR_i(q^{(t)}_{-i})$ is guaranteed by assumption of u_i and A. Then for a given $q^{(t)}$, $BR_i(q^{(t)})$ denotes a vector of actions where the player i takes an action $BR_i(q^{(t)}_{-i})$ and other players take actions $q^{(t)}_{-i}$.

Proposition 2: *Suppose that players evaluate payoff sequences by* (2). *Then, for each player i, the optimal penalty $\underline{s}^i$ satisfies the following.*

(i) $u^*_i\left(q^{(0)}(\underline{s}^i)\right) \leq u^*_i\left(BR_i(q^{(0)}(\underline{s}^i))\right) \leq \underline{v}_i$.

Proposition 2 shows that the worst equilibrium path resembles a stick-and-carrot structure derived by Abreu (1986) in the repeated Cournot oligopoly, where the players first produce a penalty output and then play the best equilibrium path from the next period onward. Here, the first period action leads to a utility level weakly lower than the utility of the optimal penalty. This implies that the utility of the continuation following $q^{(0)}(\underline{s}^i)$ is weakly higher than the utility of $q^{(0)}(\underline{s}^i)$ so that a penalty comes before the reward. Again, monotonicity, continuity, and recursivity are key properties to the above result. Note that to compare the payoff levels between $u_i\left(q^{(0)}(\underline{s}^i)\right)$ and $U_i\left(Q(\underline{s}^i)\right)$, $q^{(0)}(\underline{s}^i)$ must be evaluated by the implied utility function u^*_i, which is comparable in a scale of V_i.

On the other hand, as opposed to the Abreu' (1986) example, the stick-and-carrot structure holds only for the player's *own* optimal penalty. Thus, Proposition 2 (i) *does not* imply

$$u^*_j\left(q^{(0)}(\underline{s}^i)\right) \leq u^*_j\left(BR_j(q^{(0)}(\underline{s}^i))\right) \leq \underline{v}_j,$$

for the player j other than the player i. Because of this result, we call Proposition 2 (i) an *initial self punishment*.

4 Reverse Folk Theorem

In this section, by introducing the player whose preferences are represented by (3), we investigate how the optimal penal code changes as the players becomes more time-variability averse in the direction of the current period loss. The results in this section form a basis for defining the robustness of the Nash reversion strategy profile, which will be discussed in the next section.

4.1 Coherent Action Space

First, we introduce the player whose preferences are represented by (3).

Assumption 1: A set of players
At least, the first K players evaluate payoff sequences by the form of utility function (3), where $1 \leq K \leq I$.

By abuse of notation, K indicates both the number of players and the set of the players. We specify such K at each Proposition below. Furthermore, let $\underline{\delta}(K)$ be the vector of gain discount factors defined by $\underline{\delta}(K) \equiv (\underline{\delta}_1, ..., \underline{\delta}_K)$, and let $\overline{\delta}(K)$ be the vector of loss discount factors defined by $\overline{\delta}(K) \equiv (\overline{\delta}_1, ..., \overline{\delta}_K)$. Note that in Assumption 1, we have no restriction on the preferences for the last $I - K$ players except that the preferences are represented by (2). Thus, for some of those players, the preferences may follow (3).

For the following subsections, we impose a few restrictions on the structure of the stage game G as well as on the set of players' actions. The first assumption is concerned with the existence of the Nash equilibrium.

Assumption 2: Nash Equilibrium in the state game G
There exists a unique Nash equilibrium in the state game G, whose equilibrium action profile is denoted by $q^N = (q_1^N, q_2^N, ..., q_I^N)$.

Given Assumption 2, the *Nash reversion strategy profile* is a subgame perfect equilibrium with a form of the reversion strategy profile $s(Q, Q(s^N), \ldots, Q(s^N))$ where s^N is the strategy profile under which each player plays the static Nash equilibrium action q_i^N at every period.

For a general supergame G^∞, players can take actions freely. Because the set of equilibria can be highly complex, we restrict our attention by imposing some conditions on players' actions. For this purpose, we introduce the following assumption, which generalize a symmetric path generated by a symmetric subgame perfect equilibrium based on a symmetric stage game G.

Assumption 3: Coherent Action Space

The players' action space is restricted to A^*, called a *coherent action space*, that satisfies all of the following conditions.

(i) A^* is a closed subset of A.

(ii) $\phi_i^{-1}(A^*) = A_i$ for all i, where ϕ_i^{-1} is the projection function from A to A_i.

(iii) $q^N \in A^*$.

(iv) For any distinct $q, q' \in A^*$, if $u_i(q) > u_i(q')$ for some i, $u_j(q) > u_j(q')$ for all j.

(v) For any $q \in A^*$, if $u_i(BR_i(q)) > u_i(q)$ for some i, $u_j(BR_j(q)) > u_j(q)$ for all j.

By Assumption 3 (iii) and (v), $u_i(BR_i(q)) = u_i(q)$ if and only if $q = q^N$. It becomes clear in the proof that these conditions play a crucial role for deriving the results of this paper. Note that we do not claim that a coherent action space exists for any stage game G. We restrict our attention to the stage game G where a coherent action space exists.

We also introduce a couple of notations to define the collection of subgame perfect equilibria. Let Γ be the collection of all paths Q with $q^{(t)} \in A^*$ for all t, and let $S^*(\Gamma; \underline{\delta}(K), \overline{\delta}(K))$ be the collection of all pure strategy subgame perfect equilibria whose continuation paths on and off the equilibrium are all in Γ. This notation emphasis the dependence of the equilibria on $\underline{\delta}(K)$ and $\overline{\delta}(K)$. As shown in Lemma A.4 (see Appendix D), Proposition 1 holds for $S^*(\Gamma; \underline{\delta}(K), \overline{\delta}(K))$ replacing S^*.

4.2 The Reverse Limit of the Optimal Penal Code

In the Folk theorem, we investigate the behavior of equilibria in the limit of discount factors. In terms of a limit of discount factors, we have several choices because representation (3) has two discount factors, gain discount factor $\underline{\delta}_i$ and loss discount factor $\overline{\delta}_i$. In a usual setting of the discounted utility model, that is, (3) with $\underline{\delta}_i = \overline{\delta}_i$, we study the behavior of the equilibrium set by approaching the discount factor to one. This line of research confirms that an increase in the discount factor induces collusive behavior. As we show in Section 4.4, for representation (3), this type of behavior essentially corresponds to the case where loss discount factor $\overline{\delta}_i$ approaches to one.

Instead, representation (3) allows us to study the case where collusion becomes harder even though the players remain farsighted. Such a case is captured by letting the gain discount factor $\underline{\delta}_i$ approach to zero while keeping loss discount factor $\overline{\delta}_i$ unchanged. To see this further, if the player i's preferences are represented by (3), the utility of the optimal penal penalty $\underline{s}^i$ becomes

$$V_{i,0}(U_i(Q(\underline{s}^i))) = (1 - \underline{\delta}_i)u_i(q^{(0)}(\underline{s}^i)) + \underline{\delta}_i V_{i,1}(U_i(Q(\underline{s}^i))).$$

Note that $V_{i,1}(U_i(Q(\underline{s}^i)))$ is the utility of the payoff sequence generated by some subgame perfect equilibrium, which represents the future gain received as compensation for accepting the current penalty $u_i(q^{(0)}(\underline{s}^i))$. Because Proposition 2 implies that $V_{i,1}(U_i(Q(\underline{s}^i)))$ is weakly larger than $\underline{v}_i$, the gain factor $\underline{\delta}_i$ should be used to evaluate this sequence. As a decrease in $\underline{\delta}_i$ causes the decreased anticipation of future gain, the player i no longer accepts a sever current penalty at a given level of future compensation. Thus, the optimal penalty becomes weaker so that $\underline{v}_i$ increases. This suggests that the gain discount factor $\underline{\delta}_i$ controls the *enforceability* of the optimal penalty.

Formally, for the first K players, $\left\{\underline{\delta}_k^{(n)}\right\}$ denotes any monotonically decreasing sequence that converges to zero starting at $\underline{\delta}_k^{(0)} = \underline{\delta}_k$, where $1 \leq k \leq K$. Let $\underline{\delta}^{(n)}(K)$ be the set of gain discount factors defined by $\underline{\delta}^{(n)}(K) \equiv (\underline{\delta}_1^{(n)}, ..., \underline{\delta}_K^{(n)})$, and let $S^*(\underline{\delta}^{(n)}(K), \overline{\delta}(K))$ denote the collection of all pure strategy subgame perfect equilibria if the player k's gain discount factor is $\underline{\delta}_k^{(n)}$ for $k \in K$, while keeping other preference parameters fixed. For the first K players, we also denote by $V_{k,0}(.; \underline{\delta}'_k, \overline{\delta}'_k)$ the representation (3) based on $[\underline{\delta}'_k, \overline{\delta}'_k]$ to emphasize the dependence of the gain factor $\underline{\delta}'_k$ and the loss discount factor $\overline{\delta}'_k$.

The following proposition states the limiting behavior of the optimal penal code for a player whose preferences are represented by (3) (see Appendix C for the proof). Note that this Proposition does not require Assumptions 2 and 3.

Proposition 3: *Given Assumption 1 with a fixed $K \geq 1$, for each n, let $(\underline{s}^{1,n}, ..., \underline{s}^{I,n})$ be the optimal penal code in $S^*(\underline{\delta}^{(n)}(K), \overline{\delta}(K))$. Then for each $k \in K$, as $\underline{\delta}_k^{(n)}$ approaches to zero,*

(i) $\left|V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k\left(q^{(0)}(\underline{s}^{k,n})\right)\right|$ *converges to zero.*

(ii) $\left|u_k\left(BR_k(q^{(0)}(\underline{s}^{k,n}))\right) - u_k\left(q^{(0)}(\underline{s}^{k,n})\right)\right|$ *converges to zero.*

Proposition 2 and Proposition 3 imply that if $u_k\left(q^{(0)}(\underline{s}^{k,n})\right)$ converses to the point where $q^{(0)}(\underline{s}^{k,\infty}) \equiv \lim_{n\to\infty} q^{(0)}(\underline{s}^{k,n})$, then

$$u_k\left(q^{(0)}(\underline{s}^{k,\infty})\right) = u_k\left(BR_k(q^{(0)}(\underline{s}^{k,\infty}))\right) = V_{k,0}(U_k(Q(\underline{s}^{k,\infty})); 0, \overline{\delta}_k).$$

Here, the utility of the first period action is identical to the utility of the optimal penalty as well as the utility of the best response action. Because $u_k^*(q) = u_k(q)$ in the representation (3), this means that the utility of the optimal penalty is identical to the utility of the constant action sequence of $q^{(0)}(\underline{s}^{k,\infty})$. It is clear that if $q^{(0)}(\underline{s}^{k,\infty})$ is a Nash equilibrium action in the stage game G, $V_{k,0}(U_k(Q(\underline{s}^{k,\infty})); 0, \overline{\delta}_k)$ is the utility of the subgame perfect equilibrium where all players takes the Nash equilibrium action at any history. As we show in the following subsections, Assumption 3 becomes a workable sufficient condition for this result.

4.3 Reverse Folk Theorem

We derive an important intermediate result by investigating a limiting behavior of the optimal penal code via the convergence of the gain discount factor. As this effectively shows the convergence of the worst equilibrium, we call it the *reverse Folk theorem*, which contrast with the standard Folk theorem that is used to show the convergence of the best equilibrium.

Proposition 4: ***(Reverse Folk Theorem)*** *Given Assumption 1 with a fixed $K \geq 1$, Assumption 2, and Assumption 3, let $V_{k,t}^{(n)}(.)$ denote $V_{k,t}(.; \underline{\delta}_k^{(n)}, \overline{\delta}_k)$ for $k \in K$, and let $V_{j,t}^{(n)}(.)$ denote $V_{j,t}(.)$ for $j \in I/K$. Then, for any $s \in S^*(\Gamma; \underline{\delta}(K), \overline{\delta}(K))$, if $V_{i,0}^{(0)}(U_i(Q(s))) < V_{i,0}^{(0)}(U_i(Q(s^N)))$ for some i, then there exists n such that $s \notin S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$.*

Proposition 4 follows directly from Proposition 5 below (so that we do not provide the proof). It states that if, for some player i, the utility of an equilibrium s is less than the utility of s^N, then s is never be an optimal penalty if the first K players becomes sufficiently time-variability averse in the direction of gain. Note that in the standard Folk theorem, we set a target payoff sequence and derive the level of the discount factor that makes the target sequence an equilibrium payoff path. On the other hand, in the reverse Folk theorem, we set a target utility level of the penalty and derive the level of the gain discount factor under which the target penalty is no longer a subgame perfect equilibrium.

Note that Proposition 4 holds even for the case of $K = 1$, where there is only a single player whose preferences follow representation (3) and for whom we investigate a limiting behavior by letting a gain discount factor approach to zero.

In terms of the proof, Proposition 5 below summarizes the detailed characteristics of the limiting behavior of the optimal penal code in $S^*(\Gamma;\underline{\delta}^{(n)}(K)),\overline{\delta}(K))$ (see Appendix D for the proof).

Proposition 5: *Given Assumption 1 with a fixed $K \geq 1$, Assumption 2, and Assumption 3, let $(\underline{s}^{1,n},...,\underline{s}^{I,n})$ be the optimal penal code in $S^*(\Gamma;\underline{\delta}^{(n)}(K)),\overline{\delta}(K))$. Then, as $\underline{\delta}_k^{(n)}$ approaches to zero for all $k \in K$,*

(i) for any player $k \in K$, $q^{(0)}(\underline{s}^{k,n})$ converges to q^N.

(ii) for any player $k \in K$, $V_{k,0}(U_k(Q(\underline{s}^{i,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)$ converges to $u_k(q^N)$ for all $i \in I$.

(iii) for any player $j \in I/K$, $V_{j,0}(U_j(Q(\underline{s}^{i,n})))$ converges to $u_j^*(q^N)$ for all $i \in I/K$.

(iv) for any $s \in \bigcap_{n\geq 0} S^*(\Gamma;\underline{\delta}^{(n)}(K),\overline{\delta}(K))$, $u_i\left(q^{(0)}(s)\right) \geq u_i\left(q^N\right)$ for all $i \in I$.

(v) for any player $j \in I/K$, $Q(\underline{s}^{j,n})$ converges to $Q(s^N)$.

Proposition 5 holds when $K = 1$. Thus, it states that if at least *one* of the players' preferences are represented by (3) and if its gain discount factor $\underline{\delta}_k$ approaches to zero, the player i's utility of the optimal penalty $\underline{s}^{i,n}$ converges to the utility of the strategy profile s^N for all players i. Furthermore, except the player $k \in K$, the equilibrium path of a player's optimal penalty converge to the path induced by the strategy profile s^N where all players play the stage game Nash equilibrium action q^N at every period. As for the optimal penalty of the player $k \in K$, the initial action taken by all players converges to the stage game Nash equilibrium action.

The proof of Proposition 5 is based on Proposition 2 and Proposition 3. The sketch of the proof is as follows: First, for (i), consider $k \in K$, and let q^* be the limit of $q^{(0)}(\underline{s}^{k,n})$. Then, by the continuity of u_k, it follows from Proposition 3 (ii) that

$$u_k(q^*) = u_k(BR_k(q^*)),$$

where the coherence implied by Assumption 3 (v) leads to $q^* = q^N$. With this result, (ii) follows from Proposition 2. In terms of (iii), if the limit of $V_{j,0}(U_j(Q(\underline{s}^{j,n})))$ is smaller than $u_j^*\left(q^N\right)$, then $u_j^*\left(q^{(0)}(\underline{s}^{j,m}))\right) < u_j^*\left(q^N\right)$ after a sufficiently large m because Proposition 2 states that $u_j^*\left(q^{(0)}(\underline{s}^{j,m}))\right) \leq V_{j,0}(U_j(Q(\underline{s}^{j,m})))$. Given the coherence implied by Assumption 3 (iv), we obtain that $u_k\left(q^{(0)}(\underline{s}^{j,m}))\right) < u_k\left(q^N\right)$. Because of (ii), this relationship can be used to show that $V_{k,0}(U_k(Q(\underline{s}^{j,m}));\underline{\delta}_k^{(m)},\overline{\delta}_k) <$

$V_{k,0}(U_k(Q(\underline{s}^{k,m}));\underline{\delta}_k^{(m)},\overline{\delta}_k)$, which contradicts that $\underline{s}^{k,m}$ is the optimal penalty at m for k. Thus, $V_{j,0}(U_j(Q(\underline{s}^{j,n})))$ must converge to $u_j^*\left(q^N\right)$. A similar logic is applied to the proof of (iv). As for (v), given the above results, it follows from the strict monotonicity and continuity of W_j that $V_{j,1}(U_j(Q(\underline{s}^{j,n})))$ converges to $u_j^*\left(q^N\right)$. By repeatedly applying this argument, we show that a payoff sequence $U_j(Q(\underline{s}^{j,n}))$ converges to the sequence of $u_j^*\left(q^N\right)$. Then, (v) follows from the uniqueness of the stage game Nash equilibrium. This happens because in W_j, the importance of future payoffs relative to the current payoff is fixed for any n. On the other hand, for the player k, we can show only that $q^{(0)}(\underline{s}^{k,n})$ converges to q^N because the player becomes myopic as n increases. Lastly, (v) implies that (ii) and (iii) hold for $\underline{s}^{i,n}$ for all $i \in I/K$. We leave all of the technical details to Appendix D.

Next, we examine the conditions that guarantee the convergence of the payoff sequence $U_k(Q(\underline{s}^{k,n}))$ to $U_k(Q(s^N))$ for $k \in K$. This condition is not required for the reverse Folk theorem, but we examine it for a completeness of the argument. For this analysis, consider the path $(q^N, Q(s))$ with $s \in \bigcap_{n\geq 0} S^*(\Gamma;\underline{\delta}^{(n)}(K)),\overline{\delta}(K))$. By Proposition 4, $V_{k,0}(U_k(Q(s));\underline{\delta}_k^{(0)},\overline{\delta}_k) \geq u_k(q^N)$ so that

$$
\begin{aligned}
& V_{k,0}((u_k\left(q^N\right), U_k(Q(s)));\underline{\delta}_k^{(0)},\overline{\delta}_k) \\
&= (1-\underline{\delta}_k^{(0)})u_k(q^N) + \underline{\delta}_k^{(0)} V_{k,0}(U_k(Q(s));\underline{\delta}_k^{(0)},\overline{\delta}_k) \\
&\geq (1-\underline{\delta}_k^{(0)})u_k(BR_k(q^N)) + \underline{\delta}_k^{(0)} V_{k,0}(U_k(Q(s^N));\underline{\delta}_k^{(0)},\overline{\delta}_k) \\
&= u_k(q^N),
\end{aligned}
\tag{6}
$$

which shows that $\left((q^N, Q(s)), Q(s^N), ..., Q(s^N)\right)$ is a subgame perfect equilibrium in $S^*(\Gamma;\underline{\delta}^{(0)}(K)),\overline{\delta}(K))$. Moreover, the repeated application of Proposition 4 (treating n as 0) implies that $V_{k,0}(U_k(Q(s));\underline{\delta}_k^{(n)},\overline{\delta}_k) \geq u_k(q^N)$ for all n. Thus, (6) holds for each n so that $\left((q^N, Q(s)), Q(s^N), ..., Q(s^N)\right)$ is a subgame perfect equilibrium in $\bigcap_{n\geq 0} S^*(\Gamma;\underline{\delta}^{(n)}(K)),\overline{\delta}(K))$. Furthermore, $V_{k,0}((u_k\left(q^N\right), U_k(Q(s)));\underline{\delta}_k^{(n)},\overline{\delta}_k)$ converges to $u_k(q^N)$ as $\underline{\delta}_k^{(n)}$ approaches to zero. Therefore, for the player k, "playing q^N at time 0 and following some subgame perfect equilibrium s" becomes an optimal penalty at the limit. This leaves a possibility that $U_k(Q(\underline{s}^{k,n}))$ may not converge to $U_k(Q(s^N))$ for the player k. Note that in an incentive constraint such as (6), only the utility, not a path, of the penalty matters because $Q(\underline{s}^{k,n})$ is a path of thread, which is an out of the equilibrium path.

The following proposition shows a sufficient condition that guarantees the convergence of $U_k(Q(\underline{s}^{k,n}))$ to $U_k(Q(s^N))$, which in turn implies that the convergence of

$Q(\underline{s}^{k,n})$ to $Q(s^N)$ (see Appendix D for the proof).

Proposition 6: *Given Assumption 1 with a fixed $K \geq 1$, Assumption 2, and Assumption 3, let $(\underline{s}^{1,n}, ..., \underline{s}^{I,n})$ be the optimal penal code in $S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$. Then, for the player $k \in K$, $Q(\underline{s}^{k,n})$ converges to $Q(s^N)$ as $\underline{\delta}_k^{(n)}$ approaches to zero if A is a finite set.*

If A is a finite set, Proposition 5 (i) shows that $u_k\left(q^{(0)}(\underline{s}^{k,n})\right) = u_k\left(q^N\right)$ for a sufficiently large n. Then, Proposition 2 implies that $V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) = u_k(q^N)$. As W_k is strictly monotone and continuous, these two equalities lead to $V_{k,1}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) = u_k(q^N)$. With this result, it follows from Proposition 5 (ii), Proposition 5 (iv), and the strict monotonicity and continuity of W_k that $u_k\left(q^{(1)}(\underline{s}^{k,n})\right) = u_k\left(q^N\right)$ and $V_{k,2}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) = u_k(q^N)$. By repeatedly applying this construction, we can show that $u_k\left(q^{(t)}(\underline{s}^{k,n})\right)$ converges to $u_k\left(q^N\right)$ for all t. The conclusion follows from the uniqueness of the stage game Nash equilibrium.

4.4 Relation to the Folk Theorem

Lastly in this section, we consider an alternative limiting case where the loss discount factor $\overline{\delta}$ approaches to one. The result corresponds to the standard folk theorem of utility function (3).[10] Namely, a given level of collusion is always sustainable if the players' loss discount factor is sufficiently large. Thus, the loss discount factor $\overline{\delta}$ is a key to collusion (see Appendix E for the proof). Note that as in the case of the discounted utility model, we need to assume that all players' preferences are represented by (3).

Proposition 7: *Given Assumption 1 with $K = I$ and Assumption 2, consider a path Q, where there exists a fixed $\varepsilon > 0$ such that $u_k(q^{(t)}) - u_k\left(q^N\right) > \varepsilon$ for each k and for all t. There exists $\delta^* \in (0, 1)$ such that*

(i) if $\overline{\delta}_k \geq \delta^*$ for all k, $s(Q, Q(\underline{s}^N), ..., Q(\underline{s}^N))$ is a subgame perfect equilibrium in $S^*(\Gamma; \underline{\delta}(K)), \overline{\delta}(K))$.

Proposition 7 corresponds to the case of the discounted utility model with homogeneous or heterogenous $\overline{\delta}$, where the Nash reversion strategy profile implements Q

[10] We do not explore the full potential of the Folk theorem of this setting.

for any sufficiently large $\overline{\delta}$. In particular, Proposition 7 includes the case of the discounted utility by letting $\underline{\delta}_k = \overline{\delta}_k$. On the other hand, Proposition 7 holds regardless of $\underline{\delta}_k$, where full collusion is achieved for any sufficiently large $\overline{\delta}_k$. This result isolates the effect of the loss discount factor from the effect of the gain discount factor.

For the proof, consider the Nash reversion strategy profile $s(Q, Q(\underline{s}^N), ..., Q(\underline{s}^N))$. For this to be a subgame perfect equilibrium, for each player $k \in K$,

$$(1 - \overline{\delta}_k)u_k(BR_k(q^{(t)})) + \overline{\delta}_k V_{k,t+1}(U_k(Q(s^N))) \leq V_{k,t}(U_k(Q)) \tag{7}$$

must hold at any time t. Observe that by the assumption of Q stated in the proposition, for each k and for all t

$$u_k(BR_k(q^{(t)})) - \varepsilon > u_k(q^N),$$

and

$$V_{k,t}(U_k(Q)) - \varepsilon > V_{k,t}(U_k(Q(s^N))) = u_k(q^N),$$

which shows that $\overline{\delta}_k$ needs to be used in (7). Then for a sufficiently large $\overline{\delta}_k$, (7) is satisfied at any time t. Thus, $\overline{\delta}_k$ determines the level of *deterrence* of the penal code s^N: as $\overline{\delta}_k$ becomes larger, the player k increasingly fears the penalty s^N so that the player k avoids deviating from the collusive path. Note that the left hand side of (7) is independent of $\underline{\delta}_k$. Therefore, an increase in $\overline{\delta}_k$ opens a possibility of collusion even if the player k is nearly myopic when facing a current period loss, that is, when $\underline{\delta}_k$ is nearly zero.

5 Robustness of the Nash Reversion Strategy Profile

In this section, we show our main result, the refinement of the subgame perfect equilibria, where s^N indeed becomes a unique optimal penalty. We derive this result based on the limiting behavior of the optimal penal code

For the following argument, given $\underline{\delta}(K)$ and $\overline{\delta}(K)$, we introduce the following two sets.

$$\underline{\Delta}(\underline{\delta}(K)) \equiv \left\{ \underline{\delta}'(K) = (\underline{\delta}'_1, ..., \underline{\delta}'_K) \in (0,1)^K \, | \underline{\delta}'_k \leq \underline{\delta}_k \text{ for each } k \in K \right\}.$$

and

$$\overline{\Delta}(\overline{\delta}(K)) \equiv \left\{ \overline{\delta}'(K) = (\overline{\delta}'_1, ..., \overline{\delta}'_K) \in (0,1)^K \, \middle| \overline{\delta}'_k \geq \overline{\delta}_k \text{ for each } k \in K \right\}.$$

Thus, the players in K who evaluate payoff sequences based on $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$ and $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$ are more time-variability averse than the players in K who evaluate payoff sequences based on $\underline{\delta}(K)$ and $\overline{\delta}(K)$. We also use $\underline{\Delta}_k(\underline{\delta}(K))$ and $\overline{\Delta}_k(\overline{\delta}(K))$ to indicate the projection of $\underline{\Delta}(\underline{\delta}(K))$ and $\overline{\Delta}(\overline{\delta}(K))$ onto the kth coordinate, that is, the set of discount factors under which the kth player becomes more time-variability averse. Furthermore, let $S^*(\underline{\delta}'(K), \overline{\delta}'(K))$ be the set of all pure strategy subgame perfect equilibria of G^∞, where for each $k \in K$, the player k's gain and loss discount factors are $\underline{\delta}'_k$ and $\overline{\delta}'_k$, respectively.

We now define the set of gain/loss robust subgame perfect equilibria as follows.

Definition 1: Given Assumption 1, the set of *gain/loss robust subgame perfect equilibria* $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ is defined by

$$\begin{aligned} &S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K))) \\ &\equiv \left\{ s \in S \;\middle|\; \begin{array}{c} s \in S^*(\underline{\delta}'(K), \overline{\delta}'(K)) \\ \text{for all } \underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K)) \text{ and for all } \overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K)) \end{array} \right\}. \end{aligned}$$

The element of $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ is an equilibrium in $S^*(\underline{\delta}(K), \overline{\delta}(K))$, which is also an equilibrium under any combination of more time-variability averse players, that is an equilibrium in $S^*(\underline{\delta}'(K), \overline{\delta}'(K))$ for any $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$ and $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$. With Assumptions 2 and 3, $S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ is the gain/loss robust subgame perfect equilibria whose continuation paths on and off the equilibrium are all in Γ.

The gain/loss robust equilibrium captures the following situation:

1. At each time t and each history h^t, the preferences of the last $I - K$ players as well as the form of preferences (3) of the first K players are common knowledge.

2. For each k, the player k, as well as all other players, are unsure about the player k's degree of time-variability aversion except the upper bound of $\underline{\delta}'_k$ and the lower bound of $\overline{\delta}'_k$, that is, $\overline{\delta}_k$ and $\underline{\delta}_k$, respectively. This behavioral unsureness is common knowledge.

3. To deal with this unsureness, all players consider only the strategy that is an equilibrium under any combination of $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$ and $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$. This choice criteria is also common knowledge.

Note that we do not assume that the player k's set of discount factor changes over time. It is fixed but simply unknown. We might want to consider the set of discount factor as a type of the player so that each of the first K players has uncountable types. Furthermore, for the case of $K = 1$, we may assume more: the player $k = 1$ knows his/her own preferences, but the players other than the player $k = 1$ are unsure about the degree of time-variability aversion except the upper bound of $\underline{\delta}'_k$ and the lower bound of $\overline{\delta}'_k$, that is, $\overline{\delta}_k$ and $\underline{\delta}_k$, respectively.

Given above interpretation, we can show that at any time t and any history h^t, the set of gain/loss subgame perfect equilibria is always $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$. We then introduce a selection criteria for the optimal penalty.

Definition 2: Given Assumption 1, a penal code $(\underline{s}^{1,R}, ..., \underline{s}^{I,R})$ whose element $\underline{s}^{i,R}$ is in $S^*(\underline{\delta}(K), \overline{\delta}(K))$ is called a *gain/loss robust optimal penal code* if $(\underline{s}^{1,R}, ..., \underline{s}^{I,R})$ satisfies the following:

(i) For each player $i = 1, ..., I$, $\underline{s}^{i,R} \in S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$.

(ii) For each player $j \in I/K$,

$$V_{j,0}(U_j(Q(\underline{s}^{j,R}))) = \min \left\{ V_{j,0}(U_j(Q(s))) \,\middle|\, s \in S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K))) \right\}.$$

(iii) For each player $k \in K$ and any combination of $\overline{\delta}'_k \in \overline{\Delta}_k(\overline{\delta}(K))$ and $\underline{\delta}'_k \in \underline{\Delta}_k(\underline{\delta}(K))$,

$$V_{k,0}(U_k(Q(\underline{s}^{k,R})); \underline{\delta}'_k, \overline{\delta}'_k) = \min \left\{ V_{k,0}(U_k(Q(s)); \underline{\delta}'_k, \overline{\delta}'_k) \,\middle|\, s \in S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K))) \right\}.$$

Condition (i) states that a gain/loss robust optimal penalty itself is a gain/loss robust subgame perfect equilibrium. Condition (ii) states that for the player $j \in I/K$, a gain/loss robust optimal penalty attains the least utility among all gain/loss robust subgame perfect equilibria. Condition (iii) is essentially identical to Condition (ii), but the player k evaluates each gain/loss robust subgame perfect equilibrium via the utility function with $[\underline{\delta}'_k, \overline{\delta}'_k]$ for each $\overline{\delta}'_k \in \overline{\Delta}_k(\overline{\delta}(K))$ and $\underline{\delta}'_k \in \underline{\Delta}_k(\underline{\delta}(K))$. Thus, Condition (iii) implies that $\underline{s}^{k,R}$ is the strongest penalty regardless of the level of time-variability aversion.

Note that we cannot directly apply the definition of the optimal penal code for $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ because the value of $V_{k,0}(U_k(Q(s)); \underline{\delta}'_k, \overline{\delta}'_k)$ depends on $[\underline{\delta}'_k, \overline{\delta}'_k]$. Therefore, we need a criteria that is independent of $[\underline{\delta}'_k, \overline{\delta}'_k]$.

Now, we want to confirm the optimality of a gain/loss robust optimal penal code.

Proposition 8: ***(Optimality of a gain/loss robust optimal penal code)*** *Given Assumption 1 with a fixed $K \geq 1$, let $(\underline{s}^{1,R}, ..., \underline{s}^{I,R})$ be a gain/loss robust optimal penal code. Then for any $s \in S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$, $s(Q(s), Q(\underline{s}^{1,R}), \ldots, Q(\underline{s}^{I,R}))$ becomes a subgame perfect equilibrium in $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$.*

Proposition 8 follows directly from Definition 2. To see this, consider the case of $(\underline{\delta}(K), \overline{\delta}(K))$. It follows from Definition 2 that for any $s \in S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$, for all players i and for all t,

$$\begin{aligned} & W_i \left(u_i \left(BR_i(q^{(t)}(s))\right), V_{i,0}(U_i(Q(\underline{s}^{i,R})))\right) \\ & \leq W_i \left(u_i \left(BR_i(q^{(t)}(s))\right), V_{i,0}(U_i(Q(s')))\right) \\ & \leq V_{i,t}(U_i(Q(s))), \end{aligned} \tag{8}$$

where for the player $k \in K$, $V_{i,t}(.)$ is referred to as $V_{i,t}(.; \underline{\delta}_k, \overline{\delta}_k)$. The strategy profile s' is the subgame perfect equilibrium in $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ used in the strategy profile s as a penalty to defer the deviation.[11] Because (8) defines the relation based on the utility functions at $(\underline{\delta}(K), \overline{\delta}(K))$ and denies a profitable one deviation, Lemma A.1 implies that $s(Q(s), Q(\underline{s}^{1,R}), \ldots, Q(\underline{s}^{I,R}))$ becomes a subgame perfect equilibrium in $S^*(\underline{\delta}(K), \overline{\delta}(K))$. A similar argument also proves that $s(Q(s), Q(\underline{s}^{1,R}), \ldots, Q(\underline{s}^{I,R}))$ becomes a subgame perfect equilibrium in $S^*(\underline{\delta}'(K), \overline{\delta}'(K))$ for any $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$ and $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$.

Although it is hard to characterize all of the properties of $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$, the following hightlights the key property implied by the time-variability aversion (see Appendix F for the proof).

Proposition 9: *Suppose that players evaluate payoff sequences by (2). Given Assumption 1 with a fixed $K \geq 1$, let $(\underline{s}^{1,R}, ..., \underline{s}^{I,R})$ be a gain/loss robust optimal penal code. Then, if*

$$u_i^*(q) \geq W_i(u_i(BR_i(q)), V_{i,0}(U_i(Q(\underline{s}^{i,R})))) \tag{9}$$

is satisfied for all i, a constant path $Q = (q, q, q, ...)$ becomes an equilibrium path in $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ generated by the reversion strategy profile $s(Q, Q(\underline{s}^{1,R}), ..., Q(\underline{s}^{I,R}))$.

Proposition 9 shows that given the reversion strategy profile $s(Q, Q(s^{1,R}), ..., Q(s^{I,R}))$, any constant equilibrium path Q in $S^*(\underline{\delta}(K), \overline{\delta}(K))$ is an equilibrium path in $S^*(\underline{\delta}'(K), \overline{\delta}'(K))$ for any $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$ and any $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$. Thus, to see whether a constant

[11] By definition, $s' \in S^*(\underline{\delta}(K), \overline{\delta}(K))$. (8) is based on this result.

path Q is an equilibrium path in $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$, we only need to confirm that Q is an equilibrium path at $(\underline{\delta}(K), \overline{\delta}(K))$ under the gain/loss robust optimal penal code.

So far, we implicitly assume that $S^*(\underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ is nonempty and a gain/loss robust optimal penal code exists. However, it is nontrivial. Proposition 9 gives us a clue because it suggests that a constant path has a potential to be a gain/loss robust optimal penal code. For this, Assumptions 2 and 3 again play a crucial role.

The following is the main result of this paper (see Appendix G for the proof).

Proposition 10: ***(Optimality of the Nash Reversion Strategy Profile)*** *Given Assumption 1 with a fixed $K \geq 1$, Assumption 2, and Assumption 3, consider $S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$. Then,*

(i) $(s^N, ..., s^N)$ is a unique gain/loss robust optimal penal code.

Intuitively, Proposition 10 follows from Proposition 5: $(s^N, ..., s^N)$ is a gain/loss robust optimal penal code because for any sequence of $\left\{\underline{\delta}_1^{(n)}, ..., \underline{\delta}_k^{(n)}\right\}$,

$$S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K))) \subseteq \bigcap_{n \geq 0} S^*(\Gamma; \underline{\delta}^{(n)}(K), \overline{\delta}(K)),$$

where for $\bigcap_{n \geq 0} S^*(\Gamma; \underline{\delta}^{(n)}(K), \overline{\delta}(K))$, $(s^N, ..., s^N)$ captures the limiting behavior of the optimal penal code. A uniqueness is a consequence of Definition 2, where even at $(\underline{\delta}(K), \overline{\delta}(K))$, a gain/loss robust optimal penal code must attain the least utility. This rules out any penalty that attains the least utility only at the limit of $\left\{\underline{\delta}_1^{(n)}, ..., \underline{\delta}_k^{(n)}\right\}$, an example of which is shown in the text following Proposition 5. Again, note that Proposition 10 holds when $K = 1$. Thus, if one of the players' preferences are represented by (3) and other players are unsure about the degree of time-variability aversion, then the Nash reversion strategy profile becomes the optimal reversion strategy.

6 Summary

In this paper, we have proposed a potential answer for the empirical puzzle: the Nash reversion strategy profile, which is suboptimal in general, was used in an actual oligopoly. Our main result is that among the gain/loss robust equilibria, the Nash reversion strategy profile is the strongest reversion strategy profile. This result follows from the reverse Folk theorem, which shows the key role of the gain discount factor as a source defining the enforceability/credibility of the optimal penalty. The weakening the enforceability of the optimal penal code eventually rules out any punishment path except that induced by Nash reversion strategy profile. The reverse Folk theorem also helps us to isolate the role of the loss discount factor and confirm that it is the key to implementing collusive action because it defines a upper bound of the collusive payoff via the Nash reversion strategy profile.

7 Acknowledgment

We have benefited from comments by Atsushi Kajii, Takashi Kunimoto, and seminar participants at the Game Theory Conference 2014 at the Tokyo Institute of Technology, Keio University, Okayama University, Stony Brook International Conference on Game Theory, University of Hong Kong, Yonsei University, and World Congress of the Econometric Society 2015. Sekiguchi is thankful to Ishii Memorial Securities Research Promotion Foundation for its financial support. We gratefully acknowledge the financial support from the Japanese government in the form of research grant, Grant-in-Aid for Scientific Research (Sekiguchi, 23000001, 26245024, 26380238; Wakai, 23530219, 21K01387).

A: Proof of Proposition 1

For the proof of Proposition 1, we introduce additional notations. We denote by (s_i, s_{-i}) a strategy profile consisting of s_i and s_{-i}, where $s_{-i} \equiv (s_j)_{j \neq i}$. Let $s_t \equiv \{s_{i,t}\}_{i=1}^{I}$. We also use $q^{(t)}(s_t)$ to denote the time-t action profile $q^{(t)}(s)$ interchangeably. Moreover, for a given strategy profile s and a given history h^{t-1}, let $Q(s; h^{t-1})$ be a path

$$Q(s; h^{t-1}) = (q^{(0)}(s; h^{t-1}), q^{(1)}(s; h^{t-1}), \ldots)$$

such that (a) its play at time $0, 1, \ldots, t-1$ coincides with h^{t-1}, and (b) the players follow s from time t onward. Formally,

(a) $(q^{(0)}(s; h^{t-1}), ..., q^{(t-1)}(s; h^{t-1})) = h^{t-1}$, and

(b) $q^{(\tau)}(s; h^{t-1}) = q^{(\tau)}\left(\left\{s_{i,\tau}(q^{(0)}(s; h^{t-1}), \ldots, q^{(\tau-1)}(s; h^{t-1}))\right\}_{i=1}^{I}\right)$ for all $\tau \geq t$.

Note that $Q(s) = Q(s; h^{-1})$. For any $t > 0$, $Q(s) = Q(s; h^{t-1})$ holds if and only if $h^{t-1} = (q^{(0)}(s), ..., q^{(t-1)}(s))$.

Next, we introduce the following property.

Definition A.1 (one-deviation property): *A strategy profile s is said to satisfy the one-deviation property if for all t and all h^{t-1}, no player can increase her utility by changing her current action given the opponents' strategies and the rest of her own strategy.*

The following lemma is the key to prove Proposition 1.

Lemma A.1: *A strategy profile s is a subgame perfect equilibrium if and only if it satisfies the one-deviation property.*

Proof. The necessity of the one-deviation property follows from the definition of the subgame perfect equilibrium.

For the sufficiency, assume that a strategy profile s satisfies the one-deviation property. Suppose, by way of contradiction, that s is not a subgame perfect equilibrium. Then, there exist time t, history h^{t-1}, and a player i with a strategy s'_i such that

$$V_{i,t}(U_i(Q(s'_i, s_{-i}; h^{t-1}))) > V_{i,t}(U_i(Q(s; h^{t-1}))). \tag{10}$$

Given such s'_i, we consider a sequence of histories $\{h^\tau\}_{\tau=t}^{\infty}$ such that for all $\tau \geq t$,

$$h^\tau \equiv (q^{(0)}(s'_i, s_{-i}; h^{t-1}), \ldots, q^{(\tau)}(s'_i, s_{-i}; h^{t-1})).$$

Then, it follows from the product topology adopted to $(A)^\infty$ that $Q(s; h^\tau)$ converges to $Q(s'_i, s_{-i}; h^{t-1})$ as τ goes to infinity.

Note that $Q(s; h^t)$ is a path induced by player i's one deviation from $Q(s; h^{t-1})$. Under the path, player i switches to s'_i at time t given h^{t-1}, and then switches back to s_i from time $t+1$ onward against s_{-i}. Therefore, by the one-deviation property,

$$V_{i,t}(U_i(Q(s; h^{t-1}))) \geq V_{i,t}(U_i(Q(s; h^t))).$$

Similarly, for any $\tau > t$, $Q(s; h^\tau)$ is a path induced by player i's one deviation from $Q(s; h^{\tau-1})$. Namely, player i switches to s'_i at time τ given $h^{\tau-1}$, and then switches back to s_i from time $\tau + 1$ onward against s_{-i}. Therefore, by the one-deviation property,

$$V_{i,\tau}(U_i(Q(s; h^{\tau-1}))) \geq V_{i,\tau}(U_i(Q(s; h^\tau))).$$

By repeated application of the recursive relation (2) and the strict monotonicity of W_i in the second argument, it follows that

$$V_{i,t}(U_i(Q(s;h^{\tau-1}))) \geq V_{i,t}(U_i(Q(s;h^{\tau})))$$

for any $\tau \geq t$. It follows from iterating this relation that

$$V_{i,t}(U_i(Q(s;h^{t-1}))) \geq V_{i,t}(U_i(Q(s;h^{\tau})))$$

for any $\tau \geq t$. Because $Q(s;h^{\tau})$ converges to $Q(s_i', s_{-i}; h^{t-1})$ as $\tau \to \infty$, the continuity of $V_{i,t}$ implies

$$V_{i,t}(U_i(Q(s;h^{t-1}))) \geq V_{i,t}(U_i(Q(s_i', s_{-i}; h^{t-1}))),$$

which contradicts (10). ■

As another prerequisite for Proposition 1, we need to prove the following Lemma.

Lemma A.2: *For each i, there exist subgame perfect equilibria $\underline{s}^i$ and $\overline{s}^i$ in S^* that satisfy $V_{i,0}(U_i(Q(\underline{s}^i))) = \underline{v}_i$ and $V_{i,0}(U_i(Q(\overline{s}^i))) = \overline{v}_i$, respectively.*

For the argument below, we define paths $Q^{i,*}$ and $Q^{i,\#}$ for each i as follows: Let $\left\{Q^{i,k}\right\}_{k=1}^{\infty}$ be a sequence of subgame perfect equilibrium paths such that

$$\lim_{k\to\infty} V_{i,0}(U_i(Q^{i,k})) = \underline{v}_i.$$

Because A^∞ is a compact metric space, it has a subsequence converging to $Q^{i,*} = (q^{i,*(0)}, q^{i,*(1)}, \ldots)$ such that $V_{i,0}(U_i(Q^{i,*})) = \underline{v}_i$. Similarly, let $\left\{Q'^{i,k}\right\}_{k=1}^{\infty}$ be a sequence of subgame perfect equilibrium paths such that

$$\lim_{k\to\infty} V_{i,0}(U_i(Q'^{i,k})) = \overline{v}_i.$$

Because A^∞ is a compact metric space, it has a subsequence converging to $Q^{i,\#}$, where $V_{i,0}(U_i(Q^{i,\#})) = \overline{v}_i$.

The proofs of Lemma A.2 and Proposition 1 follow directly from Abreu (1988) with a minor modification to accommodate the change in the evaluating function. In particular, the proofs presented in Theorems 5.5 and 5.6 of Fudenberg and Tirole (1991) immediately extend to the recursive utility. The essence of their proofs is summarized by the following Lemma.

Lemma A.3: *Let* $Q^{0,*} = (q^{0,*(0)}, q^{0,*(1)}, \ldots)$ *be a path that is the limit of a sequence of subgame perfect equilibrium paths. Then, the reversion strategy profile* $s^* \equiv s(Q^{0,*}, Q^{1,*}, \ldots, Q^{I,*})$ *is a subgame perfect equilibrium.*

Proof. Suppose, by way of contradiction, s^* is not a subgame perfect equilibrium. From Lemma A.1, some player j has a profitable one deviation at some h^{t-1}, where he chooses q'_j at time t. By the construction of s^*, if no players deviate at h^{t-1} and thereafter, there exists $i \in \{0, \ldots, I\}$ and $\tau \geq 0$ such that

$$(q^{(t)}(s^*; h^{t-1}), q^{(t+1)}(s^*; h^{t-1}), \ldots) = (q^{i,*(\tau)}, q^{i,*(\tau+1)}, \ldots).$$

By this and history independence, it follows that

$$W_j(u_j(q'_j, q_{-j}^{i,*(\tau)}), \underline{v}_j) > V_{j,\tau}(U_j(Q^{i,*})).$$

Because $V_{j,\tau}$, u_j, and W_j are continuous and $Q^{i,*}$ is the limit of a sequence of subgame perfect equilibrium paths, there exists a subgame perfect equilibrium path $\hat{Q} = (\hat{q}^{(0)}, \hat{q}^{(1)}, \ldots)$ such that

$$W_j(u_j(q'_j, \hat{q}_{-j}^{(\tau)}), \underline{v}_j) > V_{j,\tau}(U_j(\hat{Q})). \tag{11}$$

Because $\hat{Q}$ is a subgame perfect equilibrium path, there exists another subgame perfect equilibrium path $\tilde{Q}$ such that

$$W_j(u_j(q'_j, \hat{q}_{-j}^{(\tau)}), V_{j,0}(U_j(\tilde{Q}))) \leq V_{j,\tau}(U_j(\hat{Q})) \tag{12}$$

holds.[12] By monotonicity of W_j, (11) and (12) imply $V_{j,0}(U_j(\tilde{Q})) < \underline{v}_j$. This is a contradiction against the definition of $\underline{v}_j$. ■

The proofs of Lemma A.2 and Proposition 1:

For Lemma A.2, the existence of $\underline{s}^i$ and $\overline{s}_i$ is a straightforward consequence of Lemma A.3 by setting $Q^{0,*} = Q^{i,*}$ and $Q^{0,*} = Q^{i,\#}$, respectively.

As for Proposition 1, fix $s^* \in S^*$. Clearly, $Q(s^*)$ is the limit of the sequence of subgame perfect equilibrium paths, $(Q(s^*), Q(s^*), \ldots)$. Because $Q(\underline{s}^i) = Q^{i,*}$ for each i, Lemma A.3 immediately proves that the reversion strategy profile $s(Q(s^*), Q(\underline{s}^1), ..., Q(\underline{s}^I))$ is a subgame perfect equilibrium. ■

[12] By the stationarity of the action space and history-independent recursive preferences, the set of subgame perfect equilibrium paths and the set of continuation paths of a subgame perfect equilibrium coincide.

B: Proof of Proposition 2

The following is the proof of Proposition 2.

Proof. Because $\underline{s}^i$ is a subgame perfect equilibrium in S^*, there exists s in S^* such that

$$v_i = V_{i,0}(U_i(Q(s))) \text{ and } \underline{v}_i = W_i(u_i(q^{(0)}(\underline{s}^i)), v_i).$$

Thus,

$$\begin{aligned}
\underline{v}_i &= W_i(u_i(q^{(0)}(\underline{s}^i)), v_i) \\
&\geq W_i(u_i(BR_i(q^{(0)}(\underline{s}^i)), \underline{v}_i) \\
&= W_i(u_i(BR_i(q^{(0)}(\underline{s}^i)), W_i(u_i(q^{(0)}(\underline{s}^i)), v^i)) \\
&\geq W_i(u_i(BR_i(q^{(0)}(\underline{s}^i)), W_i(u_i(BR_i(q^{(0)}(\underline{s}^i)), W_i(u_i(q^{(0)}(\underline{s}^i)), v^i))).
\end{aligned}$$

By repeating the above replacement, it follows from the continuity of W_i that

$$\underline{v}_i \geq u_i^*(BR_i(q^{(0)}(\underline{s}^i))).$$

As $V_{i,0}$ is strictly increasing, the conclusion follows from $u_i(BR_i(q^{(0)}(\underline{s}^i))) \geq u_i(q^{(0)}(\underline{s}^i))$. ■

C: Proof of Proposition 3

The following is the proof of Proposition 3.

Proof. For each $k \in K$, let $\left\{\underline{\delta}_k^{(n)}\right\}$ be any monotonically decreasing sequence that converges to 0 starting at $\underline{\delta}_k^{(0)} = \underline{\delta}_k$. Let $\underline{s}^{k,n}$ be the optimal penalty at $\underline{\delta}_k^{(n)}$. By (4), we denote by $\left(u_k(q^{(0)}(\underline{s}^{k,n})), V_{k,1}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k)\right)$ the utility vector satisfying

$$\begin{aligned}
&V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) \\
&= (1 - \underline{\delta}_k^{(n)}) u_k(q^{(0)}(\underline{s}^{k,n})) + \underline{\delta}_k^{(n)} V_{k,1}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k),
\end{aligned} \tag{13}$$

where by Proposition 2, the gain gain factor $\underline{\delta}_i^{(n)}$ is used to evaluate the sequence. Then (13) implies that

$$\begin{aligned}
&V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(q^{(0)}(\underline{s}^{k,n})) \\
&= \underline{\delta}_k^{(n)} \left\{ V_{k,1}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(q^{(0)}(\underline{s}^{k,n})) \right\},
\end{aligned} \tag{14}$$

where by Proposition 2,

$$0 \leq V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(q^{(0)}(\underline{s}^{k,n})).$$

Moreover, let M be defined as $\max\{|u_1^*(A)|, |u_2^*(A)|, ..., |u_I^*(A)|)\}$, which is independent of $\underline{\delta}_k^{(n)}$. Then,

$$\left| V_{k,1}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(q^{(0)}(\underline{s}^{k,n})) \right| \leq 2M.$$

Therefore, as $\underline{\delta}_k^{(n)}$ converges to zero, $\left\{ V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(q^{(0)}(\underline{s}^{k,n})) \right\}$ converges to zero, which proves (i).

As for (ii), by Proposition 2,

$$\begin{aligned} 0 &\leq V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(BR_k(q^{(0)}(\underline{s}^{k,n}))) \\ &\leq V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) - u_k(q^{(0)}(\underline{s}^{k,n})), \end{aligned}$$

which, with (14), implies (ii). ■

D : Proof of Proposition 5 and Proposition 6

We first show the following lemma.

Lemma A.4: *Suppose that players evaluate payoff sequences by the evaluating function satisfying* (2)*. Then, under Assumption 1 with a fixed $K \geq 1$, Assumption 2, and Assumption 3, Proposition 1 holds for $S^*(\Gamma; \underline{\delta}(K), \overline{\delta}(K))$ replacing S^*.*

Proof. It suffices to show that Γ is compact and $S^*(\Gamma; \underline{\delta}(K), \overline{\delta}(K))$ is nonempty.

For the first claim, it suffices to show that Γ is a closed subset of compact A^∞, which follows immediately from Assumption 3 (i). For the second claim, the strategy profile where the players choose q^N at any history is a subgame perfect equilibrium that induces a path in Γ. Thus, $S^*(\Gamma; \underline{\delta}(K), \overline{\delta}(K))$ is nonempty. ■

We provide the proofs in the following order: Proposition 5 and Proposition 6.

The proof of Proposition 5 (The following (Step 1) to (Step 8) constitute a proof).

For each $k \in K$, let $\left\{\underline{\delta}_k^{(n)}\right\}$ be a monotonically decreasing sequence that converges to zero starting at $\underline{\delta}_k^{(0)} = \underline{\delta}_k$. Let $\left(\underline{s}^{1,n}, ..., \underline{s}^{I,n}\right)$ be an optimal penal code in $S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$.

(Step 1) For each $k \in K$, as n goes to ∞, $q^{(0)}(\underline{s}^{k,n})$ converges to q^N.

Proof. For a given $k \in K$, consider a convergent subsequence $\left\{q^{(0)}(\underline{s}^{k,n_l})\right\}$ of $\left\{q^{(0)}(\underline{s}^{k,n})\right\}$, the existence of which is guaranteed by the compact metric space A. Let q^* be the limit of $\left\{q^{(0)}(\underline{s}^{k,n_l})\right\}$. Observe that

$$\begin{aligned} 0 &\leq \left|u_k\left(BR_k\left(q^{(0)}(\underline{s}^{k,n_l})\right)\right) - u_k\left(q^*\right)\right| \\ &\leq \left|u_k\left(BR_k\left(q^{(0)}(\underline{s}^{k,n_l})\right)\right) - u_k\left(q^{(0)}(\underline{s}^{k,n_l})\right)\right| + \left|u_k\left(q^{(0)}(\underline{s}^{k,n_l})\right) - u_k\left(q^*\right)\right|. \end{aligned} \tag{15}$$

Then it follows from (ii) of Proposition 3, (15), and continuity of u_j that

$$\left|u_k\left(BR_k\left(q^{(0)}(\underline{s}^{k,n_l})\right)\right) - u_k\left(q^*\right)\right| \text{ converges to zero.} \tag{16}$$

Given (16), we claim that

$$u_k\left(BR_k\left(q^*\right)\right) - u_k\left(q^*\right) = 0. \tag{17}$$

Under Assumption 3, (17) implies that for all j other than k,

$$u_j\left(BR_j\left(q^*\right)\right) - u_j\left(q^*\right) = 0,$$

which shows that

$$q^* = q^N.$$

If (17) holds, any convergent subsequence $\left\{q^{(0)}(\underline{s}^{k,n_l})\right\}$ of $\left\{q^{(0)}(\underline{s}^{k,n})\right\}$ must converge to q^N. This means that $\left\{q^{(0)}(\underline{s}^{k,n})\right\}$ converges to q^N. Therefore, it suffices to prove (17).

Suppose, by way of contradiction, that $u_k\left(BR_k\left(q^*\right)\right) - u_k\left(q^*\right) > \varepsilon > 0$ for some ε. It follows from continuity of u_k that there exists $\varepsilon' > 0$ such that for any $q' \in A$ with $\left|q' - BR_k\left(q^*\right)\right| < \varepsilon'$,

$$u_k\left(q'\right) - u_k\left(q^*\right) > \frac{\varepsilon}{2}. \tag{18}$$

Observe that

$$\begin{aligned} &\left(q_1, ..., q_{k-1}, q_k, q_{k+1}, ..., q_I\right) - BR_k\left(q^*\right) \\ &= \left(q_1, ..., q_{k-1}, q_k, q_{k+1}, ..., q_I\right) - \left(q_1^*, ..., q_{k-1}^*, BR_k(q_{-k}^*), q_{k+1}^*, ..., q_I^*\right). \end{aligned}$$

Because $q^{(0)}(\underline{s}^{k,n_l})$ converges to q^*, there exist a sufficiently large n_l and $q_k' \in A_k$ such that for all $m \geq n_l$,

$$\left|\left(q_1^{(0)}(\underline{s}^{k,m}), ..., q_{k-1}^{(0)}(\underline{s}^{k,m}), q_k', q_{k+1}^{(0)}(\underline{s}^{k,m}), ..., q_k^{(0)}(\underline{s}^{k,m})\right) - BR_k\left(q^*\right)\right| < \varepsilon' \tag{19}$$

and

$$\left|u_k\left(q^{(0)}(\underline{s}^{k,m})\right) - u_k\left(q^*\right)\right| < \frac{\varepsilon}{4}. \tag{20}$$

Then it follows from (18), (19), and (20) that for all $m \geq n_l$,

$$u_k\left(q_1^{(0)}(\underline{s}^{k,m}), ..., q_{k-1}^{(0)}(\underline{s}^{k,m}), q_k', q_{k+1}^{(0)}(\underline{s}^{k,m}), ..., q_k^{(0)}(\underline{s}^{k,m})\right) - u_k\left(q^{(0)}(\underline{s}^{k,m})\right) > \frac{\varepsilon}{4}. \tag{21}$$

Furthermore,

$$u_k\left(BR_k\left(q^{(0)}(\underline{s}^{k,m})\right)\right) \geq u_k\left(q_1^{(0)}(\underline{s}^{k,m}), ..., q_{k-1}^{(0)}(\underline{s}^{k,m}), q_k', q_{k+1}^{(0)}(\underline{s}^{k,m}), ..., q_k^{(0)}(\underline{s}^{k,m})\right). \tag{22}$$

Thus, (21) and (22) imply that for all $m \geq n_l$,

$$u_k\left(BR_k\left(q^{(0)}(\underline{s}^{k,m})\right)\right) - u_k\left(q^{(0)}(\underline{s}^{k,m})\right) > \frac{\varepsilon}{4},$$

which contradicts to (ii) of Proposition 3. This proves (17). ■

(Step 2) As n goes to ∞, $V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k)$ converges to $u_k(q^N)$.

Proof. Because $V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}_k) \leq V_{k,0}(U_k(Q(s^N)); \underline{\delta}_k^{(n)}, \overline{\delta}_k) = u_k(q^N)$ for all n, by (i) of Proposition 2, it suffices to prove that $u_k(q^{(0)}(\underline{s}^{k,n}))$ converges to $u_k(q^N)$ as n goes to ∞. Given the continuity of u_k, this follows from (Step 1). ■

For an equilibrium strategy s, consider a player j with $j \in I/K$. We denote by $\left(u_j(q^{(0)}(s)), V_{j,1}(U_j(Q(s)))\right)$ the utility vector of a player j that is an argument of W_j at time 0 satisfying (2).

(Step 3) As n goes to ∞, $V_{j,0}(U_j(Q(\underline{s}^{j,n})))$ converges to $u_j^*(q^N) = V_{j,0}(U_j(Q(s^N)))$ for a player $j \in I/K$.

Proof. Suppose, by way of contradiction, that there exist a player j with $j \in I/K$, $\varepsilon > 0$, and a subsequence $\{V_{j,0}(U_j(Q(\underline{s}^{j,n_l})))\}$ of $\{V_{j,0}(U_j(Q(\underline{s}^{j,n})))\}$ such that for all n_l,

$$\left|V_{j,0}(U_j(Q(\underline{s}^{j,n_l}))) - u_j^*(q^N)\right| \geq \varepsilon. \tag{23}$$

Given that $V_{j,0}(U_j(Q(\underline{s}^{j,n_l}))) \leq u_j^*(q^N) = V_{j,0}(U_j(Q(s^N)))$, it follows form (23) and Proposition 2 that for all n_l,

$$u_j^*(q^{(0)}(\underline{s}^{j,n_l})) \leq V_{j,0}(U_j(Q(\underline{s}^{j,n_l}))) \leq u_j^*(q^N) - \varepsilon. \tag{24}$$

Let $\underline{A}^*$ be the set of actions that satisfies

$$\underline{A}^* \equiv \left\{ q \in A^* \; u_j^*(q) \leq u_j^*(q^N) - \varepsilon \right\}.$$

By definition, $q^{(0)}(\underline{s}^{j,n_l}) \in \underline{A}^*$ for all n_l. Also, by continuity of u_j^*,

$$u_j^*(\overline{\underline{A}^*}) \subset \overline{u_j^*(\underline{A}^*)}.$$

Thus, there exists $q^* \in \overline{\underline{A}^*}$ such that

$$q^* \equiv \arg\max_{q \in \overline{\underline{A}^*}} \left\{ u_j^*(q) \right\}. \tag{25}$$

By construction,

$$u_j^*(q^*) < u_j^*(q^N).$$

Furthermore, for some $k \in K$, the continuity of u_k implies that

$$u_k(\overline{\underline{A}^*}) \subset \overline{u_k(\underline{A}^*)}.$$

Let $q^{**} \in \overline{\underline{A}^*}$ be defined by

$$q^{**} \equiv \arg\max_{q \in \overline{\underline{A}^*}} \left\{ u_k(q) \right\}.$$

By Assumption 3, if $u_k(q^{**}) = u_k^*(q^N)$, then $u_j^*(q^{**}) = u_j^*(q^N)$, which contradicts (25). Therefore, there exists $\varepsilon' > 0$ such that

$$u_k(q^{**}) < u_k^*(q^N) - \varepsilon'. \tag{26}$$

Given (26), it follows from Proposition 2 that there exists a sufficiently large n_l such that for all $m \geq n_l$,

$$\begin{aligned}
&V_{k,0}(U_k(Q(\underline{s}^{j,m})); \underline{\delta}_k^{(m)}, \overline{\delta}_k) \\
&= \min_{\delta_k \in [\underline{\delta}_k^{(m)}, \overline{\delta}_k]} \left[(1-\delta_k) u_k(q^{(0)}(\underline{s}^{j,m})) + \delta_k V_{k,1}(U_k(Q(\underline{s}^{j,m})); \underline{\delta}_k^{(m)}, , \overline{\delta}_k) \right] \\
&\leq (1 - \underline{\delta}_k^{(m)}) u_k(q^{(0)}(\underline{s}^{j,m})) + \underline{\delta}_k^{(m)} M \\
&< u_k\left(q^N\right) - \frac{\varepsilon'}{2}
\end{aligned}$$

where M is a bound for the absolute value of utility of subgame perfect equilibria. However, (Step 2) shows that for a sufficiently large $n_l' > n_l$,

$$0 \leq u_k(q^N) - V_{k,0}(U_k(Q(\underline{s}^{k,n_l'}); \underline{\delta}_k^{(n_l')}, \overline{\delta}_k) < \frac{\varepsilon'}{4}.$$

The above two inequalities imply that

$$V_{k,0}(U_k(Q(\underline{s}^{j,n'_l})); \underline{\delta}_k^{(n'_l)}, \overline{\delta}_k) < V_{k,0}(U_k(Q(\underline{s}^{k,n'_l}); \underline{\delta}_k^{(n'_l)}, \overline{\delta}_k) - \frac{\varepsilon'}{4}$$

which contradicts that $\underline{s}^{k,n'_l}$ is the optimal penalty at $\underline{\delta}_k^{(n'_l)}$. ■

(Step 4) For any $s \in S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$ where $u_j\left(q^{(0)}(s)\right) < u_j\left(q^N\right)$ for some j, there exists $n^* > n$ such that $s \notin S^*(\Gamma; \underline{\delta}^{(n^*)}(K)), \overline{\delta}(K))$.

Note that (Step 4) essentially proves Proposition 5 (iv).

Proof. For any $s \in S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$ where $u_j\left(q^{(0)}(s)\right) < u_j\left(q^N\right)$ for some j, there exists $\varepsilon > 0$ such that

$$0 < \varepsilon < \left|u_j\left(q^{(0)}(s)\right) - u_j\left(q^N\right)\right|.$$

By Assumption 3, there exist a player $k \in K$ and $\varepsilon' > 0$ such that

$$u_k\left(q^{(0)}(s)\right) < u_k\left(q^N\right) \text{ and } 0 < \varepsilon' < \left|u_k\left(q^{(0)}(s)\right) - u_k\left(q^N\right)\right|. \tag{27}$$

It follows from Proposition 2 that there exists a sufficiently large n' such that for all $m \geq n'$,

$$\begin{aligned}
&V_{k,0}(U_k(Q(s)); \underline{\delta}_k^{(m)}, \overline{\delta}_k) \\
&= \min_{\delta_k \in [\underline{\delta}_k^{(m)}, \overline{\delta}_k]} \left[(1-\delta_k)u_k(q^{(0)}(s)) + \delta_k V_{k,1}(U_k(Q(s)); \underline{\delta}_k^{(m)}, \overline{\delta}_k)\right] \\
&\leq (1-\underline{\delta}_k^{(m)})u_k(q^{(0)}(s)) + \underline{\delta}_k^{(m)}\overline{u_k(A)} \\
&< u_k\left(q^N\right) - \frac{\varepsilon'}{2}
\end{aligned}$$

where $\overline{u_k(A)}$ is a upper bound of $u_k(A)$. However, (Step 2) shows that for a sufficiently large $n^* > n'$,

$$0 \leq u_k(q^N) - V_{k,0}(U_k(Q(\underline{s}^{k,n^*}); \underline{\delta}_k^{(n^*)}, \overline{\delta}_k) < \frac{\varepsilon'}{4}.$$

The above two inequalities imply that

$$V_{k,0}(U_k(Q(s)); \underline{\delta}_k^{(n^*)}, \overline{\delta}_k) < V_{k,0}(U_k(Q(\underline{s}^{k,n^*})); \underline{\delta}_k^{(n^*)}, \overline{\delta}_k) - \frac{\varepsilon'}{4}.$$

Because $\underline{s}^{k,n^*}$ is the worst subgame perfect equilibrium for the player k, this shows that $s \notin S^*(\Gamma; \underline{\delta}^{(n^*)}(K)), \overline{\delta}(K))$. ■

(Step 5) For a player $j \in I/K$, as n goes to ∞, $U_j(Q(\underline{s}^{j,n}))$ converges to $U_j(Q(s^N))$.

Proof. For a player $j \in I/K$, let $\left(u_j\left(q^{(0)}(\underline{s}^{j,n})\right), V_{j,1}(U_j(Q(\underline{s}^{j,n})))\right)$ be the utility vector of the player j that is an argument of W_j at time 0 satisfying (2). It follows from (Step 4) that

$$\liminf_{n\to\infty} u_j\left(q^{(0)}(\underline{s}^{j,n})\right) \geq u_j\left(q^N\right). \tag{28}$$

Moreover, because $V_{j,1}(U_j(Q(\underline{s}^{j,n})))$ is the utility of some subgame perfect equilibrium in $S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$,

$$V_{j,1}(U_j(Q(\underline{s}^{j,n}))) \geq V_{j,0}(U_j(Q(\underline{s}^{j,n}))). \tag{29}$$

Given (28) and (29), it follows from (Step 3) and the strict monotonicity of W_j that

$$\lim_{n\to\infty} u_j\left(q^{(0)}(\underline{s}^{j,n})\right) = u_j(q^N) \text{ and } \lim_{n\to\infty} V_{j,1}(U_j(Q(\underline{s}^{j,n}))) = u_j^*(q^N).$$

Next, let $\left(u_j\left(q^{(1)}(\underline{s}^{j,n})\right), V_{j,2}(U_i(Q(\underline{s}^{j,n})))\right)$ be the utility vector of the player j that is an argument of W_j at time 1 satisfying (2). Because $V_{j,1}(U_j(Q(\underline{s}^{j,n})))$ is the utility of some subgame perfect equilibrium in $S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$, it follows from (Step 4) and the strict monotonicity of W_j that (28) replacing 0 with 1 and (29) replacing 1 with 2 hold. Given this result, it follows from (Step 3) and the strict monotonicity of W_j that

$$\lim_{n\to\infty} u_j\left(q^{(1)}(\underline{s}^{j,n})\right) = u_j(q^N) \text{ and } \lim_{n\to\infty} V_{j,2}(U_j(Q(\underline{s}^{j,n}))) = u_j^*(q^N).$$

By repeating the above construction, we obtain that for all t, $u_j\left(q^{(t)}(\underline{s}^{j,n})\right)$ converges to $u_j\left(q^N\right)$, which means, in the product topology generated by the metric on A, that $U_j(Q(\underline{s}^{j,n}))$ converges to $U_j(Q(s^N))$. ■

(Step 6) For a player $j \in I/K$, as $\underline{\delta}_k^{(n)}$ approaches to zero for all $k \in K$, $Q(\underline{s}^{j,n})$ converges to $Q(s^N)$ for all player j.

Proof. We prove this claim by showing that for each t, $q^{(t)}(\underline{s}^{j,n})$ converges to q^N. Suppose, by way of contradiction, that for some t and $j \in I/K$, there exist an open set $B \subseteq A$ that contains q^N and a convergent subsequence $\left\{q^{(t)}(\underline{s}^{j,n_l})\right\}$ of $\left\{q^{(t)}(\underline{s}^{j,n})\right\}$ such that $q^{(t)}(\underline{s}^{j,n_l}) \notin B$ for all n_l. Because the stage game Nash equilibrium is unique, it follows from Assumption 3 and the argument leading to (17) that there exists $\varepsilon > 0$ such that

$$u_j\left(BR_j(q^{(t)}(\underline{s}^{j,n_l}))\right) - u_j\left(q^{(t)}(\underline{s}^{j,n_l})\right) > \varepsilon > 0 \text{ for all } n_l. \tag{30}$$

(Otherwise, $\lim_{n\to\infty}\left|u_j\left(BR_j(q^{(t)}(\underline{s}^{j,n_l}))\right)-u_j\left(q^{(t)}(\underline{s}^{j,n_l})\right)\right|=0$. Then the argument leading to (17) shows that $u_k\left(BR_k\left(q^*\right)\right)-u_k\left(q^*\right)=0$, where q^* is the limit of $\left\{q^{(0)}(\underline{s}^{k,n_l})\right\}$. By Assumption 3, this leads to $q^*=q^N$.)

Furthermore, by (Step 5), it follows from the strictly monotonicity and continuity of V_j that as $\underline{\delta}_k^{(n)}$ approaches to zero,

$$V_{j,t+1}(U_j(Q(\underline{s}^{j,n_l})))\text{ converges to }u_j^*\left(q^N\right). \tag{31}$$

Given that W_j is strictly monotone and continuous, it follows from (Step 3), (30), and (31) that for a sufficiently large n_l,

$$\begin{aligned}&W_j\left(u_j\left(BR_j(q^{(t)}(\underline{s}^{j,n_l}))\right),V_{j,0}(U_j(Q(\underline{s}^{j,n_l})))\right)\\&>W_j\left(u_j\left(q^{(t)}(\underline{s}^{j,n_l}))\right),V_{j,t+1}(U_j(Q(\underline{s}^{j,n_l})))\right)\\&=V_{j,t}(U_j(Q(\underline{s}^{j,n_l}))).\end{aligned} \tag{32}$$

However, (32) implies that the player j can increase the utility by deviating from $\underline{s}^{j,n_l}$ at time t and accepts the optimal penalty $\underline{s}^{j,n_l}$. This contradicts that $\underline{s}^{j,n_l}$ is a subgame perfect equilibrium in $S^*(\Gamma;\underline{\delta}^{(n_l)}(K)),\overline{\delta}(K))$. ■

(Step 7) For a player $k\in K$, as n goes to ∞, $V_{k,0}(U_k(Q(\underline{s}^{i,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)$ converges to $u_k(q^N)$ for all $i\in I$.

Proof. For a given k, for each $k'\in K$,

$$\begin{aligned}V_{k,0}(U_k(Q(\underline{s}^{k,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)&\le V_{k,0}(U_k(Q(\underline{s}^{k',n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)\\&\le(1-\underline{\delta}_k^{(n)})u_k(q^{(0)}(\underline{s}^{k',n}))+\underline{\delta}_k^{(n)}M,\end{aligned} \tag{33}$$

where M is the upper bound of $V_{k,0}(S^*(\Gamma;\underline{\delta}^{(n)}(K)),\overline{\delta}(K)))$. As n increases, it follows from (Step 1) that

$$u_k(q^{(0)}(\underline{s}^{k',n}))\text{ converges to }u_k(q^N). \tag{34}$$

Thus, (Step 2), (33), and (34) imply that

$$V_{k,0}(U_k(Q(\underline{s}^{k',n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)\text{ converges to }u_k(q^N).$$

As for $j\in I/K$, it follows from the continuity of $V_{k,0}$ and (Step 6) that

$$V_{k,0}(U_k(Q(\underline{s}^{j,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)\text{ converges to }u_k(q^N).$$

The above two results proves the claim. ■

(Step 8) For a player $j \in I/K$, as n goes to ∞, $V_{j,0}(U_k(Q(\underline{s}^{i,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)$ converges to $u_k(q^N)$ for all $i \in I/K$.

Proof. This follows from the continuity of $V_{j,0}$ and (Step 6). ■

The proof of Proposition 6.

For the following, let $k \in K$ be fixed.

(Step 9) As $\underline{\delta}_k^{(n)}$ approaches to zero, $u_k\left(q^{(t)}(\underline{s}^{k,n})\right)$ and $V_{k,t}(U_k(Q(\underline{s}^{k,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)$ converge to $u_k\left(q^N\right)$ for all t.

Proof. Because A is a finite set and $s^N \in S^*(\Gamma;\underline{\delta}^{(m)}(K)),\overline{\delta}(K))$ for all m, it follows from Proposition 2 and (Step 4) that there exists n_1 such that for all $m \geq n_1$, $u_k\left(q^{(0)}(\underline{s}^{k,m})\right) = u_k\left(q^N\right)$ and $V_{k,0}(U_k(Q(\underline{s}^{k,m}));\underline{\delta}_k^{(m)},\overline{\delta}_k) = u_k(q^N)$. Because W_k is strictly monotone and continuous, this implies that $V_{k,1}(U_k(Q(\underline{s}^{k,m}));\underline{\delta}_k^{(m)},\overline{\delta}_k) = u_k(q^N)$. Again, because A is a finite set, (Step 4) implies that there exists n_2 satisfying $n_2 \geq n_1$ such that for all $m \geq n_2$, $u_k\left(q^{(1)}(\underline{s}^{k,m})\right) \geq u_k\left(q^N\right)$. Given that $V_{k,1}(U_k(Q(\underline{s}^{k,m}));\underline{\delta}_k^{(m)}) = u_k(q^N)$, it then follows from (Step 2) and strict monotonicity and continuity of W_k that $u_k\left(q^{(1)}(\underline{s}^{k,m'})\right) = u_k\left(q^N\right)$ and $V_{i,2}(U_i(Q(\underline{s}^{i,m'}));\underline{\delta}_i^{(m')},\overline{\delta}_k) = u_j(q^N)$ for a sufficiently large m'. By repeatedly applying the same construction, as $\underline{\delta}_k^{(n)}$ approaches to zero, $u_k\left(q^{(t)}(\underline{s}^{k,n})\right)$ and $V_{k,t}(U_k(Q(\underline{s}^{k,n}));\underline{\delta}_k^{(n)},\overline{\delta}_k)$ converge to $u_k\left(q^N\right)$ for all t. ■

(Step 10) As $\underline{\delta}_k^{(n)}$ approaches to zero, $Q(\underline{s}^{k,n})$ converges to $Q(s^N)$.

Proof. Given (Step 9), this follows from (Step 6) by replacing $\underline{s}^{j,n}$ with $\underline{s}^{k,n}$, where (32) is replaced by the following

$$\begin{aligned}
&\min_{\delta_k \in [\underline{\delta}_k^{(n_l)},\overline{\delta}_k]} \left[(1-\delta_k)u_k\left(BR_k(q^{(t)}(\underline{s}^{k,n_l}))\right) + \delta_k V_{k,0}(U_k(Q(\underline{s}^{k,n_l}));\underline{\delta}_k^{(n_l)},\overline{\delta}_k)\right] \\
&= (1-\delta_k^*)u_k\left(BR_k(q^{(t)}(\underline{s}^{k,n_l}))\right) + \delta_k^* V_{k,0}(U_k(Q(\underline{s}^{k,n_l}));\underline{\delta}_k^{(n_l)},\overline{\delta}_k) \\
&> (1-\delta_k^*)u_k\left(q^{(t)}(\underline{s}^{k,n_l})\right) + \delta_k^* V_{k,t+1}(U_k(Q(\underline{s}^{k,n_l}));\underline{\delta}_k^{(n_l)},\overline{\delta}_k) \\
&\geq \min_{\delta_k \in [\underline{\delta}_k^{(n_l)},\overline{\delta}_k]} \left[(1-\delta_k)u_k\left(q^{(t)}(\underline{s}^{k,n_l})\right) + \delta_k V_{k,t+1}(U_k(Q(\underline{s}^{k,n_l}));\underline{\delta}_k^{(n_l)},\overline{\delta}_k)\right].
\end{aligned}$$

Here, δ_k^* is the effective selection of the discount factor for the minimization of the first line. This contradicts that $\underline{s}^{k,n_l}$ is a subgame perfect equilibrium in $S^*(\Gamma;\underline{\delta}^{(n_l)}(K)),\overline{\delta}(K))$.
■

E : Proof of Proposition 7

Proof. Consider the Nash reversion strategy profile $s(Q, Q(\underline{s}^N), ..., Q(\underline{s}^N))$, where there exists a fixed $\varepsilon > 0$ such that $u_i(q_i^{(t)}) - u_i\left(q^N\right) > \varepsilon$ for each i and for all t. Then, it follows from representation (3) that for each i,

$$V_{i,t}(U_i(Q)) - \varepsilon > V_{i,t}(U_i(Q(s^N))) = u_i(q^N) \tag{35}$$

at any time t. Given that $u_i(A)$ is bounded, it follows from (35) that for each i, there exists δ_i^* satisfying $0 < \delta_i^* < 1$ such that at any time t,

$$(1 - \delta_i^*)u_i(BR_i(q^{(t)})) + \delta_i^* V_{i,0}(U_i(Q(s^N))) < V_{i,t}(U_i(Q)).$$

Let δ^* be defined by $\delta^* \equiv \max\{\delta_1^*, ..., \delta_I^*\}$. Then, for each i and for all $\overline{\delta}_i$ satisfying $\delta^* \leq \overline{\delta}_i < 1$,

$$(1 - \overline{\delta}_i)u_i(BR_i(q^{(t)})) + \overline{\delta}_i V_{i,0}(U_i(Q(s^N))) \leq V_{i,t}(U_i(Q)) \tag{36}$$

at any time t because

$$u_i(BR_i(q^{(t)})) \geq u_i(q_i^{(t)}) > u_i\left(q^N\right) = V_{i,0}(U_i(Q(s^N)))$$

shows that $\overline{\delta}_i$ must be used in (36). Thus, if $\overline{\delta}_i \geq \delta^*$ for all i, it follows from (36) that $s(Q, Q(\underline{s}^N), ..., Q(\underline{s}^N))$ is a subgame perfect equilibrium in S^*. ■

F : Proof of Proposition 9

Proof. For a player $j \in I/K$, (9) shows that the play j has no incentive to deviate from the constant path of q. As for a player $k \in K$, (9) shows that

$$\begin{aligned} V_{k,0}(U_k(Q)) &= u_k(q) \\ &\geq W_k(u_k(BR_k(q)), V_{k,0}(U_i(Q(\underline{s}^{k,R})))) \\ &= (1 - \overline{\delta}_k)u_k(BR_k(q)) + \overline{\delta}_k V_{k,0}(U_k(Q(\underline{s}^{k,R}))), \end{aligned}$$

where Q is the constant sequence of q. Here, $\overline{\delta}_k$ is used to evaluate in the last line because W_k is strictly monotone and continuous and $u_k(BR_k(q)) \geq u_k(q)$, so that $u_k(BR_k(q^{(t)})) \geq V_{k,0}(U_k(Q(\underline{s}^{k,R})))$. The above inequality implies that for any $\overline{\delta}'_k \in \overline{\Delta}_k(\overline{\delta}(K))$ and any $\underline{\delta}'_k \in \underline{\Delta}_k(\underline{\delta}(K))$,

$$\begin{aligned} u_k(q) &\geq (1 - \overline{\delta}'_k)u_k(BR_k(q)) + \overline{\delta}'_k V_{k,0}(U_k(Q(\underline{s}^{k,R}))) \\ &\geq (1 - \overline{\delta}'_k)u_k(BR_k(q)) + \overline{\delta}'_k V_{k,0}(U_k(Q(\underline{s}^{k,R})); \underline{\delta}'_k, \overline{\delta}'_k). \end{aligned}$$

This shows that the play k has no incentive to deviate from the constant path of q at any any $\overline{\delta}'_k \in \overline{\Delta}_k(\overline{\delta}(K))$ and any $\underline{\delta}'_k \in \underline{\Delta}_k(\underline{\delta}(K))$. ■

G : Proof of Proposition 10

Proof. First, Definition 2 (i) follows because $s^N \in S^*(\Gamma; \underline{\delta}'(K), \overline{\delta}'(K))$ for any $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$ and any $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$.

Second, for the first K players, $\left\{\underline{\delta}_k^{(n)}\right\}$ denotes any monotonically decreasing sequence that converges to zero starting at $\underline{\delta}_k^{(0)} = \underline{\delta}_k$, where $1 \leq k \leq K$. Given Definition 2 (i), Proposition 5 (iii) implies that for $j \in I/K$,

$$V_{j,0}(U_j(Q(s^N))) = \min \left\{ V_{j,0}(U_j(Q(s))) \,\middle|\, s \in \bigcap_n S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K)) \right\}.$$

Definition 2 (ii) follows because $S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K))) \subseteq \bigcap_n S^*(\Gamma; \underline{\delta}^{(n)}(K)), \overline{\delta}(K))$ and $s^N \in S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$.

Third, for Definition 2 (iii), suppose, by way of contradiction, that there exist $s \in S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ and $(\underline{\delta}'(K), \overline{\delta}'(K))$ satisfying $\underline{\delta}'(K) \in \underline{\Delta}(\underline{\delta}(K))$ and $\overline{\delta}'(K) \in \overline{\Delta}(\overline{\delta}(K))$ such that for some $k \in K$,

$$u_k\left(q^N\right) = V_{k,0}(U_k(Q(s^N)); \underline{\delta}'_k, \overline{\delta}'_k) > V_{k,0}(U_k(Q(s)); \underline{\delta}'_k, \overline{\delta}'_k). \tag{37}$$

For the first K players, $\left\{\underline{\delta}_k^{(n)}\right\}$ denotes any monotonically decreasing sequence that converges to zero starting at $\underline{\delta}_k^{(0)} = \underline{\delta}'_k$, where $1 \leq k \leq K$. Then by (3), for any $n \geq 0$,

$$V_{k,0}(U_k(Q(s)); \underline{\delta}_k^{(n)}, \overline{\delta}'_k) \geq V_{k,0}(U_k(Q(s)); \underline{\delta}_k^{(n+1)}, \overline{\delta}'_k). \tag{38}$$

Furthermore, it follows from $s \in S^*(\Gamma; \underline{\delta}^{(n)}(K), \overline{\delta}'(K))$ that for the optimal penalty $\underline{s}^{k,n} \in S^*(\Gamma; \underline{\delta}^{(n)}(K), \overline{\delta}'(K))$,

$$V_{k,0}(U_k(Q(s)); \underline{\delta}_k^{(n)}, \overline{\delta}'_k) \geq V_{k,0}(U_k(Q(\underline{s}^{k,n})); \underline{\delta}_k^{(n)}, \overline{\delta}'_k). \tag{39}$$

Given (38) and (39), Proposition 5 (ii) implies

$$V_{k,0}(U_k(Q(s)); \underline{\delta}'_k, \overline{\delta}'_k) \geq u_k\left(q^N\right),$$

which contradicts (37). Then Definition 2 (iii) follows from $s^N \in S^*(\Gamma; \underline{\Delta}(\underline{\delta}(K)), \overline{\Delta}(\overline{\delta}(K)))$ and $u_k\left(q^N\right) = V_{k,0}(U_k(Q(s^N)); \underline{\delta}'_k, \overline{\delta}'_k)$ for any $\overline{\delta}'_k \in \overline{\Delta}_k(\overline{\delta}(K))$ and $\underline{\delta}'_k \in \underline{\Delta}_k(\underline{\delta}(K))$.

As for the uniqueness of a gain/loss robust optimal penal code $(s^N, ..., s^N)$, let $(\underline{s}^{1,R}, ..., \underline{s}^{I,R})$ be a gain/loss robust optimal penal code. For the following proof, for each $k \in K$, we use $V_{k,t}(U_k(Q(s)))$ to denote $V_{k,t}(U_k(Q(s)); \underline{\delta}_k, \overline{\delta}_k)$, which is equivalent to $W_k\left(u_k\left(q^{(t)}(s)\right), V_{i,t+1}(U_k(Q(s)); \underline{\delta}_k, \overline{\delta}_k)\right)$ that is strictly monotone in both arguments.

First, we claim that for each $\underline{s}^{i,R}$, $u_i\left(q^{(t)}(\underline{s}^{i,R})\right) = u_i\left(q^N\right)$ for all t. Because W_i is strictly monotone in both arguments, Conditions (ii) and (iii) of Definition 2 and Proposition 5 (iv) imply that

$$u_i\left(q^{(0)}(\underline{s}^{i,R})\right) \geq u_i\left(q^N\right) \text{ and } V_{i,1}(U_i(Q(\underline{s}^{i,R}))) \geq u_i^*\left(q^N\right). \tag{40}$$

Given that $\underline{s}^{i,R}$ is an optimal penalty,

$$V_{i,0}(U_i(Q(\underline{s}^{i,R}))) = W_i\left(u_i\left(q^{(0)}(\underline{s}^{i,R})\right), V_{i,1}(U_i(Q(\underline{s}^{i,R})))\right) = u_i^*\left(q^N\right). \tag{41}$$

Because W_i is strictly monotone in both arguments, (41) and (40) imply that

$$u_i\left(q^{(0)}(\underline{s}^{i,R})\right) = u_i\left(q^N\right) \text{ and } V_{i,1}(U_i(Q(\underline{s}^{i,R}))) = u_i^*\left(q^N\right).$$

By repeatedly applying the above construction, for all t,

$$u_i\left(q^{(t)}(\underline{s}^{i,R})\right) = u_i\left(q^N\right) \text{ and } V_{i,t+1}(U_i(Q(\underline{s}^{i,R}))) = u_i^*\left(q^N\right), \tag{42}$$

which proves the claim.

Now, we claim that $q^{(t)}(\underline{s}^{i,R}) = q^N$ for all t. Suppose, by way of contradiction, that for some t, $q^{(t)}(\underline{s}^{i,R}) \neq q^N$. Because the stage game Nash equilibrium is unique, Assumption 3 implies that there exists $\varepsilon > 0$ such that for *any* player j,

$$u_j\left(BR_j(q^{(t)}(\underline{s}^{i,R}))\right) - u_j\left(q^{(t)}(\underline{s}^{i,R})\right) > \varepsilon > 0.$$

Given (42), it follows from the strict monotonicity and continuity of W_i that

$$\begin{aligned} &W_i\left(u_i\left(BR_i(q^{(t)}(\underline{s}^{i,R}))\right), V_{i,0}(U_i(Q(\underline{s}^{i,R})))\right) \\ &= W_i\left(u_i\left(BR_i(q^{(t)}(\underline{s}^{i,R}))\right), u_i^*\left(q^N\right)\right) \\ &> W_i\left(u_i\left(q^{(t)}(\underline{s}^{i,R})\right), u_i^*\left(q^N\right)\right) \\ &= W_i\left(u_i\left(q^{(t)}(\underline{s}^{i,R})\right), V_{i,t+1}(U_i(Q(\underline{s}^{i,R})))\right). \end{aligned} \tag{43}$$

However, (43) implies that the player i can increase the utility by deviating from $\underline{s}^{i,R}$ at time t and accepts the optimal penalty $\underline{s}^{i,R}$. This contradicts that $\underline{s}^{i,R}$ is a subgame perfect equilibrium in $S^*(\Gamma; \underline{\delta}(K)), \overline{\delta}(K))$. ■

References

[1] Abreu, D. (1986): "Extremal Equilibria of Oligopolistic Supergames," *Journal of Economic Theory*, Vol. 39, pp. 191—225.

[2] Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, Vol. 56, pp. 383—396.

[3] Chade, H., P. Prokopovych, and L. Smith (2008): "Repeated Games with Present-biased Preferences," *Journal of Economic Theory*, Vol. 139, pp. 157—175.

[4] Epstein, L. (1983): "Stationary Cardinal Utility and Optimal Growth under Uncertainty," *Journal of Economic Theory*, Vol. 31, pp. 133—152.

[5] Epstein, L. and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, Vol. 57, pp. 937—969.

[6] Frederick, S., G. Loewenstein, and T. O'Donoghue (2002): "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, Vol. 40, pp. 351—401.

[7] Fudenberg, D., and J. Tirole (1991): *Game Theory*, Cambridge, The MIT Press.

[8] Igami, M. and T. Sugaya (2022): "Measuring the Incentive to Collude: The Vitamin Cartels. 1990-1999," *The Review of Economic Studies*, Vol. 89(3), pp.1460-1494.

[9] Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Vol. 47, pp. 263—292.

[10] Kochov, A., and Y. Song (2023): "Intertemporal Hedging and Trade in Repeated Games with Recursive Utility," *Econometrica*, Vol. 91(6), pp. 2333-2369.

[11] Koopmans, T. (1960): "Stationary Ordinal Utility and Impatience," *Econometrica*, Vol. 28, pp. 287—309.

[12] Loewenstein, G., and D. Prelec (1992): "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *The Quarterly Journal of Economics*, Vol. 107, pp. 573—597.

[13] Obara, I., and J. Park (2014): "Repeated Games with General Time Preference," Mimeo.

[14] Obara, I., and J. Park (2017): "Repeated Games with General Discounting," *Journal of Economic Theory*, Vol. 172, pp. 348—375.

[15] Peleg, B., and M. Yaari (1973) "On the Existence of a Consistent Course of Action When Tastes are Changing," *Review of Economic Studies*, Vol. 40, pp. 391—401.

[16] Thaler, R. (1981): "Some Empirical Evidence on Dynamic Inconsistency," *Economic Letters*, Vol. 8, pp. 201—207.

[17] Sekiguchi, T. and K. Wakai (2016): "Repeated Games with Recursive Utility: Cournot Duopoly under Gain/Loss Asymmetry," Discussion Paper No. E-16-006, Graduate School of Economics, Kyoto University.

[18] Shalev, J. (1997): "Loss Aversion in a Multiple-period Model," *Mathematical Social Sciences,* Vol. 33, pp. 203—226.

[19] Tversky, A., and D. Kahneman (1991): "Loss Aversion in Riskless Choice: A Reference-dependent Model," *The Quarterly Journal of Economics*, Vol. 106, pp. 1039-1061.

[20] Uzawa, H. (1968): "Time Preference, the Consumption Function, and Optimum Asset Holdings," In J. Wolfe, editor, *Value, Capital and Growth: Papers in Honour of Sir John Hicks*, Aldine, Chicago.

[21] Wakai, K. (2008): "A Model of Utility Smoothing," *Econometrica*, Vol. 76, 137—153.