

# Repeated Games with General Discounting\*

Ichiro Obara

University of California, Los Angeles

Jaeok Park

Yonsei University

August 7, 2015

## Abstract

We introduce a general class of time discounting, which includes time-inconsistent ones, into repeated games with perfect monitoring. A strategy profile is called an agent subgame perfect equilibrium if there is no profitable one-shot deviation at any history. We characterize strongly symmetric agent subgame perfect equilibria for repeated games with symmetric stage game. We find that the harshest punishment takes different forms given different biases. When players are future biased, the harshest punishment is supported by a version of stick-and-carrot strategy. When players are present biased, the harshest punishment may take a more complex form. In particular, the worst punishment path may need to be cyclical. We also find that the worst punishment payoff is different from the stage game minmax payoff even when players are patient. For some class of discounting, we show that the worst punishment payoff is larger than the stage game minmax payoff with present bias and smaller than the stage game minmax payoff with future bias. We also characterize the set of limit equilibrium payoffs as the length of periods converges to 0, without changing the intertemporal structure of biases.

*JEL Classification:* C73, D03

*Keywords:* Hyperbolic Discounting, Present Bias, Repeated Game, Time Inconsistency.

## 1 Introduction

Repeated game is a very useful tool to analyze cooperation/collusion in dynamic environments. It has been heavily applied in different areas in economics such as industrial organization, dynamic macroeconomics, and international trade etc. The central feature of repeated games is that incentive is provided intertemporally and endogenously. Thus the time preference plays a crucial role for the analysis of repeated games. However, almost all the

---

\*Obara gratefully acknowledges support from National Science Foundation grant SES-1135711. Park gratefully acknowledges financial support from the Yonsei University Future-leading Research Initiative of 2014 and the hospitality of the Department of Economics at UCLA during his visits.

works assume a particular type of time preference: discounted sum of payoffs/utilities with geometric discounting. Although geometric discounting is both reasonable and tractable, it is important to explore other types of discounting to understand which behavioral features depend on the assumption of geometric discounting and which does not.<sup>1</sup>

In this paper, we introduce a general class of time discounting, which includes time-inconsistent ones, into repeated games with perfect monitoring. Our formulation includes geometric discounting, quasi-hyperbolic discounting, and genuine hyperbolic discounting as a special case.<sup>2</sup>

A strategy profile is called an agent subgame perfect equilibrium (ASPE) if there is no profitable one-shot deviation at any history. We characterize strongly symmetric agent subgame perfect equilibria for repeated games with symmetric stage game. The symmetric stage game we use is standard and includes the Cournot/Bertrand competition game as a special case.

We focus on two types of discounting: one with *present bias* and one with *future bias*. A player exhibits present bias if a current self puts relatively more weight on payoffs in the near future than any future self. Conversely a player exhibits future bias if a current self puts relatively more weight on payoffs in the far future than any future self. Since a sequence of discount factors can be identified as a probability distribution, we will borrow the terminology of stochastic ordering (such as stochastic dominance etc.) to define different versions of those biases precisely.

The following is a partial list of our main results.

1. With general discounting, the best equilibrium path is still a stationary one.
2. When the players are future biased, the worst punishment path is given by a version of stick-and-carrot path where there are multiple periods with “stick” before the players enjoy “carrot”.
3. When the players are present biased, the worst punishment path takes a more complex form. Through a recursive characterization of ASPE paths, we show that the worst punishment path may exhibit a nonstationary, cyclical pattern. If we restrict attention to an equilibrium with “forgiving punishment” where the stage game payoffs are weakly increasing along any punishment path, then the worst punishment path in this class is given by a stick-and-carrot path where the gap between and stick and

---

<sup>1</sup>See Frederick, Loewenstein and O’Donoghue [4] for a critical review of a variety of models of time preferences.

<sup>2</sup>Quasi-hyperbolic discounting or so called  $\beta$ - $\delta$  discounting has been used in so many papers, for example, Laibson [6, 7] and O’Donoghue and Rabin [9] to name a few. Chade, Prokopovych and Smith [3] is the first to apply quasi-hyperbolic discounting in the context of repeated games as far as we know.

carrot is relatively narrow and the carrot may need to be a suboptimal stationary path unlike the standard case.

4. For a certain class of preferences, the worst punishment payoff is lower than the stage game minmax payoff with future-bias and higher than the stage game minmax payoff with present-bias. In particular, we obtain a closed form solution of the worst punishment payoff with  $\beta$ - $\delta$  discounting as  $\delta \rightarrow 1$  and show that it is (strictly) increasing in  $\beta$ .
5. We characterize the limit equilibrium payoff set for the case with future bias as the length of periods goes to 0 (i.e. as the players play the stage game more and more frequently), without changing the intertemporal structure of biases.

Regarding the last result, note that this exercise is impossible in the standard  $\beta$ - $\delta$  discounting framework. When the length of periods becomes shorter, presumably the bias parameter  $\beta$  would change. In addition, the assumption that the bias disappears after one period would be problematic. Our framework allows us to define an intertemporal bias structure independent of the length of periods. More specifically, we fix a bias structure in continuous time, then analyze repeated games in discrete time.

### Related Literature

A seminal work in dynamic oligopoly is Abreu [1], who studies worst punishments and best equilibrium paths of repeated symmetric games with geometric discounting. This work aims to extend Abreu's work by allowing for non-geometric discounting. Chade, Prokopovych and Smith [3] provides a recursive characterization of agent subgame perfect equilibria with quasi-hyperbolic discounting in the style of Abreu, Pearce and Stacchetti [2]. Kochov and Song [5] studies repeated games with endogenous discounting (which still induces a time-consistent preference) and shows that the players must cooperate eventually in any efficient equilibrium for the repeated prisoners' dilemma game.

We introduce the model in the next section. We examine the structure of the best equilibrium in Section 3. We study the structure of the worst punishment with future bias in Section 4 and the one with present bias in Section 5. In Section 6, we provide a characterization of limit equilibrium payoffs for the case with future bias as the length of periods converges to 0.

## 2 The Model

### 2.1 Repeated Symmetric Games with General Discounting

Our model is the standard model of repeated (symmetric) game with perfect monitoring, except that we allow for a relatively more general class of time discounting. We first introduce the symmetric stage game.

Let  $N = \{1, \dots, n\}$  be the set of  $n$  players. Each player's action set is given by  $A = [0, a^*]$  for some  $a^* > 0$ . Players' payoff functions are symmetric and given by the same continuous function  $\tilde{\pi} : A^n \rightarrow \mathbb{R}$ . Player  $i$ 's payoff is  $\tilde{\pi}(a_i, a_{-i})$  when player  $i$  chooses  $a_i$  and the other players choose  $a_{-i}$  or any permutation of it. Together they define a symmetric stage game  $G = (N, A, \tilde{\pi})$ .

The time is discrete and denoted by  $t = 1, 2, \dots$ . In each period, players choose actions simultaneously given the complete knowledge of past actions. A history at the beginning of period  $t$  is  $h_t = (a^1, \dots, a^{t-1}) \in H_t := (A^n)^{t-1}$  for  $t \geq 2$ . Let  $H = \bigcup_{t=1}^{\infty} H_t$  with  $H_1 := \{\emptyset\}$ . We only consider pure strategies. Thus player  $i$ 's strategy  $s_i$  is a mapping from  $H$  to  $A$ . Let  $S$  be the set of player  $i$ 's strategies. For any  $s_i \in S$ ,  $s_i|_h \in S$  is player  $i$ 's continuation strategy after  $h \in H$ .

Every player uses the same discounting operator  $V(\pi^1, \pi^2, \dots) \in \mathbb{R}$  to evaluate a sequence of stage-game payoffs  $(\pi^1, \pi^2, \dots) \in \mathbb{R}^\infty$ . We assume that this (linear) operator takes the following form:

$$V(\pi^1, \pi^2, \dots) = \sum_{t=1}^{\infty} \delta(t) \pi^t,$$

where  $\delta(t) > 0$  for all  $t$  and  $\sum_{t=1}^{\infty} \delta(t) = 1$  (by normalization).<sup>3</sup> Clearly the standard geometric discounting is a special case where  $\delta(t) = (1 - \delta)\delta^{t-1}$ . This operator is used at any point of time to evaluate a sequence of payoffs in the future. For example, a sequence of payoffs  $(\pi^\tau, \pi^{\tau+1}, \dots)$  from period  $\tau$  is worth  $\sum_{t=1}^{\infty} \delta(t) \pi^{\tau+t-1}$  if evaluated in period  $\tau$ . But the same sequence of payoffs is worth  $\sum_{t=\tau}^{\infty} \delta(t) \pi^t$  if evaluated in period 1. Of course time inconsistency would arise in general for such time preferences. Let  $\Gamma$  be the set of all such discounting operators.

A strategy profile  $s$  generates a sequence of action profiles  $a(s) \in (A^n)^\infty$  and a sequence of payoffs for player  $i$ :  $\pi_i(s) = (\tilde{\pi}(a_i^1(s), a_{-i}^1(s)), \tilde{\pi}(a_i^2(s), a_{-i}^2(s)), \dots)$ . The value of this sequence generated by  $s$  to player  $i$  is denoted simply by  $V_i(a(s))$ . Stage game  $G$  and discount factors  $\{\delta(t)\}_{t=1}^{\infty}$  define a repeated game, which is denoted by  $(G, \{\delta(t)\}_{t=1}^{\infty})$ .

We focus on equilibria where every player chooses the same action after any history, i.e.

---

<sup>3</sup>Note that  $V$  satisfies "continuity at infinity." Hence some preferences such as limit-of-averages preferences are excluded.

we study *strongly symmetric equilibrium*, which we will just call equilibrium from now on. Then it is useful to introduce the following notations. Let  $\pi(a)$  be each player's payoff when every player plays  $a \in A$ . Let  $\pi^*(a) = \max_{a_i \in A} \tilde{\pi}(a_i, a, \dots, a)$  be the maximum payoff that player  $i$  can achieve when every player other than  $i$  plays  $a \in A$ . Note that  $\pi^*(a) \geq \pi(a)$  for all  $a \in A$  and that  $a$  is a symmetric Nash equilibrium of the stage game if and only if  $\pi(a) = \pi^*(a)$ . The *symmetric minmax payoff* is defined by  $\min_{a \in A} \pi^*(a)$ , which we normalize to 0.

We make the following regularity assumptions on  $\pi(a)$  and  $\pi^*(a)$  throughout the paper.

**Assumption 1.**  $\pi$  is increasing with  $\pi(a^*) > 0$ .  $\pi^*$  is nondecreasing and increasing when  $\pi^*(a) > 0$ . Both  $\pi$  and  $\pi^*$  are differentiable at any  $a$  where  $\pi(a) > 0$ . There exists the unique symmetric Nash equilibrium  $a^{NE} \in (0, a^*)$ , where  $\pi(a^{NE}) > 0$  and  $\pi'(a^{NE}) > 0$ .  $\pi^*(a) - \pi(a)$  is nonincreasing when  $a < a^{NE}$  and nondecreasing when  $a > a^{NE}$ .<sup>4</sup>

Typical graphs of the functions  $\pi(a)$  and  $\pi^*(a)$  are illustrated in Figure 1. Differentiability is assumed just for the sake of exposition. Most of our results do not depend on differentiability. Since  $\pi$  is increasing,  $\pi(a^*)$  is the unique action that achieves the maximum symmetric payoff. We will use  $\bar{\pi} = \pi(a^*)$  and  $\underline{\pi} = \pi(0)$  to denote the maximum and the minimum stage game payoffs, respectively. Since  $\pi^*$  is nondecreasing and the symmetric minmax payoff is 0,  $\pi^*(0) = 0$ . Note that  $\underline{\pi} = \pi(0) < \pi^*(0) = 0$  as 0 is not a Nash equilibrium. The last part of the assumption means that the gain from a unilateral deviation becomes larger (or unchanged) as the action profile moves away from the unique symmetric Nash equilibrium.

All these assumptions are standard and satisfied in many applications. We show that a Cournot competition game with capacity constraint satisfies the above assumptions.

### Example: Cournot Game

Consider a Cournot game where the inverse demand function is given by  $p(Q) = \max\{A - BQ, 0\}$  for some  $A, B > 0$ , where  $Q = \sum_{i=1}^n q_i$ . Each firm's marginal cost is constant  $c \in (0, A)$ . Assume that each firm can produce up to the capacity constraint  $\bar{q}$ .<sup>5</sup> We assume that  $\bar{q}$  is large enough to satisfy  $A - Bn\bar{q} - c < 0$ . Then  $\pi$  and  $\pi^*$  are given as

<sup>4</sup>We use increasing/decreasing and nondecreasing/nonincreasing to distinguish strict monotonicity and weak one.

<sup>5</sup>The existence of such a bound on the production set can be assumed without loss of generality in the standard case (Abreu [1]). However this is a substantial assumption in our setting as the equilibrium notion we use is not even the standard Nash equilibrium. See footnote 9.

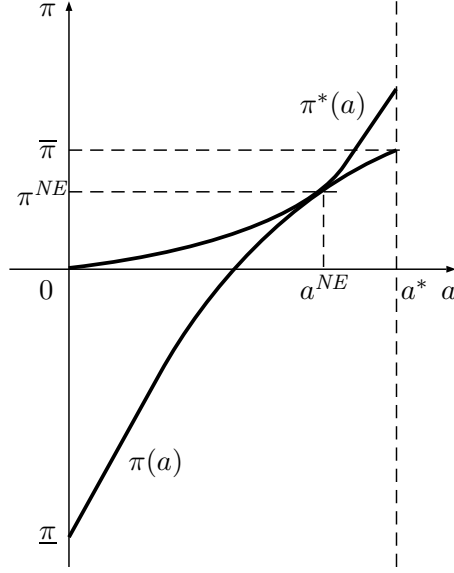


Figure 1: Graphs of  $\pi(a)$  and  $\pi^*(a)$ .

follows.

$$\pi(q) = (\max\{A - Bnq, 0\} - c)q,$$

$$\pi^*(q) = \frac{1}{B} \left( \max \left\{ \frac{A - B(n-1)q - c}{2}, 0 \right\} \right)^2.$$

Let  $q^* = \frac{A-c}{2Bn}$  be the production level that maximizes the joint profit. For each  $q \in (0, q^*)$ , there exists  $\tilde{q} \in (q^*, \bar{q})$  such that  $\pi(\tilde{q}) = \pi(q)$  and  $\pi^*(\tilde{q}) < \pi^*(q)$ . This means that we can replace  $q$  with  $\tilde{q}$  in any equilibrium. So it is without loss of generality to restrict attention to the production set  $[q^*, \bar{q}]$ . Identify  $\bar{q}$  with 0 and  $q^*$  with  $a^*$ . Then  $\pi$  and  $\pi^*$  satisfies all the assumptions except for the last one about  $\pi^*(a) - \pi(a)$ . This assumption is satisfied if  $(n+1)c > \frac{n-1}{n}A$  (then  $\pi^*(q) - \pi(q)$  is increasing for  $q > q^{NE}$ ).

## 2.2 Present Bias and Future Bias

As already mentioned, when a player evaluates a path  $a^\infty \in A^\infty$  that begins in period  $\tau$  in the beginning of the game, he uses discount factors  $(\delta(\tau), \delta(\tau+1), \dots)$ . We can normalize them so that their sum is equal to 1. For  $\tau = 1, 2, \dots$  and  $t = 1, 2, \dots$ , we define

$$\delta_\tau(t) = \frac{\delta(t + \tau - 1)}{1 - \sum_{k=1}^{\tau-1} \delta(k)}.$$

Let  $c_\tau(a^\infty) = \sum_{t=1}^{\infty} \delta_\tau(t)\pi(a^t)$  for  $\tau = 1, 2, \dots$ .  $c_\tau$  is the payoff function that is used to evaluate paths starting in period  $\tau$  in the beginning of the game. Note that  $\delta_1(t) = \delta(t)$  for all  $t$  and thus  $c_1(a^\infty) = V(a^\infty)$ , where  $V(a^\infty) = V(\pi(a^1), \pi(a^2), \dots)$  denotes the value of the path  $a^\infty \in A^\infty$ .

We say that a player may exhibit “present bias” or “future bias” when  $\{\delta_\tau(t)\}$  and  $\{\delta_{\tau'}(t)\}$  are systematically different unlike the standard case with geometric discounting. Since discount factors sum up to 1, they can be regarded as probability distributions. So we use the language of stochastic ordering to define a variety of biases.

**Definition 1** (Present Bias and Future Bias via MLRP).

1. **(Present Bias via MLRP):**  $\{\delta'(t)\}_{t=1}^{\infty} \in \Gamma$  is more present biased relative to  $\{\delta''(t)\}_{t=1}^{\infty} \in \Gamma$  in the MLRP (monotone likelihood ratio property) sense if  $\frac{\delta'(t_1)}{\delta'(t_2)} \geq \frac{\delta''(t_1)}{\delta''(t_2)}$  for all  $t_1$  and  $t_2$  such that  $t_1 < t_2$ .  $\{\delta(t)\}_{t=1}^{\infty} \in \Gamma$  exhibits present bias in the MLRP sense if  $\{\delta_\tau(t)\}_{t=1}^{\infty} \in \Gamma$  is more present biased relative to  $\{\delta_{\tau'}(t)\}_{t=1}^{\infty} \in \Gamma$  in the MLRP sense for any  $\tau$  and  $\tau'$  such that  $\tau < \tau'$ .
2. **(Future Bias via MLRP):**  $\{\delta'(t)\}_{t=1}^{\infty} \in \Gamma$  is more future biased relative to  $\{\delta''(t)\}_{t=1}^{\infty} \in \Gamma$  in the MLRP sense if  $\frac{\delta'(t_2)}{\delta'(t_1)} \geq \frac{\delta''(t_2)}{\delta''(t_1)}$  for all  $t_1$  and  $t_2$  such that  $t_1 < t_2$ .  $\{\delta(t)\}_{t=1}^{\infty} \in \Gamma$  exhibits future bias in the MLRP sense if  $\{\delta_\tau(t)\}_{t=1}^{\infty} \in \Gamma$  is more future biased relative to  $\{\delta_{\tau'}(t)\}_{t=1}^{\infty} \in \Gamma$  in the MLRP sense for any  $\tau$  and  $\tau'$  such that  $\tau < \tau'$ .

To see that this definition is consistent with the standard notion of biases, consider the following example. When comparing period 2 payoff  $\pi^2$  and period 3 payoff  $\pi^3$ , “period 2-self” puts relatively more weight on  $\pi^2$  than period 1-self if a player is present biased. This means that  $\frac{\delta(1)}{\delta(2)} \geq \frac{\delta(2)}{\delta(3)}$ . This is equivalent to  $\frac{\delta(1)}{\delta(2)} \geq \frac{\delta_2(1)}{\delta_2(2)}$ , which corresponds to the case of  $\tau = 1, \tau' = 2, t_1 = 1, t_2 = 2$  in the above definition of present bias.<sup>6</sup>

The next definition of biases is based on the notion of first-order stochastic dominance.

**Definition 2** (Present Bias and Future Bias via FOSD).

1. **(Present Bias via FOSD):**  $\{\delta'(t)\}_{t=1}^{\infty} \in \Gamma$  is more present biased relative to  $\{\delta''(t)\}_{t=1}^{\infty} \in \Gamma$  in the FOSD sense if  $\sum_{t=1}^k \delta'(t) \geq \sum_{t=1}^k \delta''(t)$  for all  $k = 1, 2, \dots$ .  $\{\delta(t)\}_{t=1}^{\infty} \in \Gamma$  exhibits present bias in the FOSD sense if  $\{\delta_\tau(t)\}_{t=1}^{\infty} \in \Gamma$  is more present biased relative to  $\{\delta_{\tau'}(t)\}_{t=1}^{\infty} \in \Gamma$  in the FOSD sense for any  $\tau$  and  $\tau'$  such that  $\tau < \tau'$ .
2. **(Future Bias via FOSD):**  $\{\delta'(t)\}_{t=1}^{\infty} \in \Gamma$  is more future biased relative to  $\{\delta''(t)\}_{t=1}^{\infty} \in \Gamma$  in the FOSD sense if  $\sum_{t=1}^k \delta''(t) \geq \sum_{t=1}^k \delta'(t)$  for all  $k = 1, 2, \dots$ .  $\{\delta(t)\}_{t=1}^{\infty} \in \Gamma$  ex-

<sup>6</sup>If  $\{\delta(t)\}_{t=1}^{\infty}$  is more present biased than  $\{\delta_k(t)\}_{t=1}^{\infty}$  in the MLRP sense for any  $k$ , then  $\{\delta_\tau(t)\}_{t=1}^{\infty}$  is more present biased than  $\{\delta_{\tau+k}(t)\}_{t=1}^{\infty}$  in the MLRP sense for any  $\tau$  and any  $k$ . Hence it suffices to set  $\tau = 1$  in Definition 1.

hibits future bias in the FOSD sense if  $\{\delta_\tau(t)\}_{t=1}^\infty \in \Gamma$  is more future biased relative to  $\{\delta_{\tau'}(t)\}_{t=1}^\infty \in \Gamma$  in the FOSD sense for any  $\tau$  and  $\tau'$  such that  $\tau < \tau'$ .

This notion of present bias captures the effect that the tendency to put relatively more weight on period  $t$  payoff against the continuation payoff from period  $t + 1$  diminishes as  $t$  increases. Clearly, if a player is present biased (resp. future biased) in the MLRP sense, then the player is present biased (resp. future biased) in the FOSD sense.

Note that all these biases are defined with weak inequalities. Hence any of these biased preferences include the standard geometric discounting as a special case. A player exhibits strict present/future bias in the MLRP/FOSD sense if all the inequalities in the corresponding definition hold strictly. For example,  $\{\delta'(t)\}_{t=1}^\infty \in \Gamma$  is strictly more present biased relative to  $\{\delta''(t)\}_{t=1}^\infty \in \Gamma$  in the MLRP sense, if  $\frac{\delta'(t_1)}{\delta'(t_2)} > \frac{\delta''(t_1)}{\delta''(t_2)}$  for all  $t_1$  and  $t_2$  such that  $t_1 < t_2$ .

We often use following class of discounting with  $\delta \in (0, 1)$  and  $\beta_t > 0$ ,  $t = 0, 1, \dots$ , which is not normalized:

$$V(a^\infty) = \beta_0 \pi(a^1) + \beta_1 \delta \pi(a^2) + \beta_2 \delta^2 \pi(a^3) + \beta_3 \delta^3 \pi(a^4) + \dots$$

We assume that  $\beta_0 = 1$  without loss of generality. We also assume that  $\beta_t$ ,  $t = 1, 2, \dots$ , is bounded above, which implies that the weighted sum of stage game payoffs is finite given any  $\delta \in (0, 1)$ . When  $\beta_t$ ,  $t = 0, 1, \dots$ , satisfies these assumptions, we call this discounting  $\{\beta_t\}_t$ -weighted discounting. This is a useful class of discounting as it nests many familiar discountings. It reduces to the standard geometric discounting if  $\beta_t = 1$  for all  $t$ . It corresponds to the standard quasi-hyperbolic ( $\beta$ - $\delta$ ) discounting (with present bias) if  $\beta_1 = \beta_2 = \beta_3 = \dots = \beta$  for some  $\beta \in (0, 1)$ . We can also represent a genuine hyperbolic discounting by choosing appropriate  $\beta_t$ ,  $t = 1, 2, \dots$ . It is straightforward to show that  $\{\beta_t\}_t$ -weighted discounting exhibits strict present bias (resp. future bias) in the MLRP sense if and only if  $\frac{\beta_{t+1}}{\beta_t}$  is increasing (resp. decreasing) with respect to  $t$ , in which case the ratio converges in  $[0, 1]$  as  $t$  goes to infinity.

### 2.3 Agent Subgame Perfect Equilibrium

In standard subgame perfect equilibrium, each player can use any possible continuation strategy at any history as a deviation. Without time consistency, however, it is not obvious which type of deviations a player would think possible, since it may be the case that some deviation is profitable from the viewpoint of the current self, but not from the viewpoint of future selves.

In this paper, we assume that each player believes that future selves (as well as all the other players) would stick to the equilibrium strategy if he deviates in the current period.



One interpretation of this assumption is that a player is a collection of agents who share the same preference and act only once. This type of equilibrium notion is standard and used in numerous papers such as Phelps and Pollak [11] and Peleg and Yaari [10]. We adapt this notion of equilibrium and call it *Agent Subgame Perfect Equilibrium (ASPE)*. The formal definition is as follows.

**Definition 3** (Agent Subgame Perfect Equilibrium). A strategy profile  $s \in S$  is an agent subgame perfect equilibrium (ASPE) if

$$V_i(a(s|h_t)) \geq V_i((a'_i, a_{-i}^1(s_{-i}|h_t)), a(s|h_t, (a'_i, a_{-i}^1(s_{-i}|h_t))))$$

for any  $a'_i \in A$ , any  $h_t \in H$  and any  $i \in N$ .

To understand this definition, compare it to the standard definition of subgame perfect equilibrium.

**Definition 4** (Subgame Perfect Equilibrium). A strategy profile  $s \in S$  is a subgame perfect equilibrium (SPE) if

$$V_i(a(s|h_t)) \geq V_i(a(s'_i, s_{-i}|h_t))$$

for any  $s'_i \in S$ , any  $h_t \in H$  and any  $i \in N$ .

All possible deviations are considered for SPE, whereas only one-shot deviations are considered for ASPE. With the standard geometric discounting, the one-shot deviation principle applies: if there is no profitable one-shot deviation, then there is no profitable deviation. Hence there is no gap between ASPE and SPE. With general discounting, the one-shot deviation principle would fail because of time inconsistency. Therefore the set of SPE is in general a strict subset of ASPE.<sup>7,8,9</sup>

Agent subgame perfect equilibrium  $s^*$  is strongly symmetric if every player plays the same action at any history, i.e.  $s_i^*(h_t) = s_j^*(h_t)$  for any  $h_t \in H$  and any  $i, j \in N$ . There

---

<sup>7</sup>In general, the set of equilibria would shrink as more deviations are allowed. However ASPE remains unchanged even if we allow for a larger class of finite period deviations where every deviating-self along a deviation path is required to gain strictly more than what he would have obtained by playing according to the continuation equilibrium. See Obara and Park [8].

<sup>8</sup>Chade, Prokopovych and Smith [3] shows that the sets of equilibrium *payoffs* for these two notions of equilibrium coincide for  $\beta$ - $\delta$  discounting.

<sup>9</sup>Unbounded stage game payoffs, which we exclude by assumption, would be another source of gap between SPE and ASPE. For example, we would have many “crazy” ASPE even with geometric discounting if there is no capacity constraint (hence the symmetric payoff  $\pi$  is unbounded below) in the Cournot game. In particular, we can construct the following type of ASPE with an arbitrary low payoff: every firm produces some large  $q'$  and makes negative profits because otherwise every firm would produce even larger amount  $q'' \gg q'$  in the next period and so on. Such behavior cannot be sustained in SPE, or even in Nash equilibrium if  $q'$  is too large, but can be sustained in ASPE. So the existence of capacity constraint  $\bar{q}$  is a substantial assumption that excludes this type of ASPE.

exists at least one strongly symmetric agent subgame perfect equilibrium in  $(G, \{\delta(t)\}_{t=1}^\infty)$  for any  $\{\delta(t)\}_{t=1}^\infty$  because we assume that there exists a symmetric Nash equilibrium for  $G$  and playing this symmetric equilibrium at any history is an ASPE.

Let  $A^{ASPE} \subseteq A^\infty$  be the set of (strongly symmetric) agent subgame perfect equilibrium paths. Since the time preference is invariant, the set of ASPE continuation paths from any period is also  $A^{ASPE}$ . We define  $\bar{v}_\tau = \sup_{a^\infty \in A^{ASPE}} c_\tau(a^\infty)$  and  $\underline{v}_\tau = \inf_{a^\infty \in A^{ASPE}} c_\tau(a^\infty)$  for  $\tau = 1, 2, \dots$ . Obara and Park [8] shows that  $\bar{v}_\tau$  and  $\underline{v}_\tau$  can be exactly attained (i.e. sup and inf can be replaced by max and min) for any  $\tau$  in more general settings. We denote any equilibrium path to achieve  $\bar{v}_\tau$  and  $\underline{v}_\tau$  by  $\bar{a}_\tau^\infty$  and  $\underline{a}_\tau^\infty$  respectively. Of particular interest among these values is  $\underline{v}_2$ , which is the the worst equilibrium continuation payoff from period 2, since it determines the size of maximum punishment. We will simply denote it by  $\underline{v}$  and call it the *worst punishment payoff*. We also call  $\underline{a}_2^\infty$  a *worst punishment path* and simply denote it by  $\underline{a}^\infty$ . As in the usual case, we can always use  $\underline{v}$  as the punishment payoff without loss of generality when we check incentive constraints.

### 3 Best Equilibrium

In the standard case, the best (strongly symmetric) equilibrium payoff is always achieved by a constant path. We first observe that this result holds even with general discounting.

The following proposition shows that the best equilibrium path  $\bar{a}_\tau^\infty$  to achieve  $\bar{v}_\tau$  is given by the unique constant path independent of  $\tau$ .

**Proposition 1.** *For all  $\tau = 1, 2, \dots$ , a best equilibrium (continuation) path  $\bar{a}_\tau^\infty$  is unique and given by a constant path  $(a^\circ, a^\circ, \dots)$  where  $a^\circ \in (a^{NE}, a^*]$  is independent of  $\tau$ . Moreover,  $a^\circ = a^*$  when  $\delta(1)$  is small enough.*

The proof of this proposition is standard. Let  $a^\circ$  be the most efficient symmetric action profile that can be achieved in any equilibrium. Then the constant path  $(a^\circ, a^\circ, \dots)$  must be an equilibrium path (hence the best equilibrium path) as well. The incentive constraint to play  $a^\circ$  is satisfied because the value of this path is at least as high as the value of any equilibrium path where  $a^\circ$  is played in the initial period and we can always use the same worst punishment path.

This proposition also says that the best equilibrium payoff is strictly higher than the stage game Nash equilibrium payoff. This follows from differentiability of  $\pi$  and  $\pi^*$ . The one-shot deviation constraint for any constant path  $(a, a, \dots)$  is given by

$$\delta(1) (\pi^*(a) - \pi(a)) \leq (1 - \delta(1)) (\pi(a) - \underline{v}).$$

Note that  $a = a^{NE}$  satisfies the above constraint and  $(\pi^*)'(a^{NE}) - \pi'(a^{NE}) = 0$ ; hence the

deviation gain on the left-hand side is roughly 0 around  $a^{NE}$  as there is no first-order effect. On the other hand, the continuation payoff loss on the right-hand side increases strictly at  $a^{NE}$ . Hence this constraint is slack for  $a$  that is slightly above  $a^{NE}$ . For any  $a > a^{NE}$ , the above incentive constraint is satisfied for sufficiently small  $\delta(1)$ , because  $\underline{v} \leq \pi(a^{NE})$ . Hence we can achieve the most efficient action  $a^*$  and payoff  $\bar{\pi}$  at ASPE when  $\delta(1)$  is small enough.

Since  $\bar{v}_\tau = \pi(a^o)$  for all  $\tau$ , we will denote this common value by  $\bar{v}$ . To compute  $\bar{v}$ , we need to know how harsh a punishment can be in ASPE. In the following sections, we study the structure of worst punishment paths and characterize the set of all equilibrium payoffs under various assumptions.

## 4 Characterization of Equilibrium: Future Bias

The structure of worst punishment paths and the set of equilibrium payoffs depend on the structure of intertemporal bias. It turns out that the case of future bias is simpler and more tractable. Thus we first examine the case of future bias.

As a preliminary, we prove the following property of the worst punishment path that holds in general, not just for future bias.

**Proposition 2.** *For all  $\tau = 1, 2, \dots$ ,  $\underline{v}_\tau < \pi(a^{NE})$  and  $\underline{a}_\tau^1 \leq a^{NE}$  for any worst punishment path  $\underline{a}_\tau^\infty$ .*

The first part of the proposition follows from differentiability of the payoff functions as was the case with Proposition 1. In particular, we can show that  $(a', a^{NE}, a^{NE}, \dots)$  can be an equilibrium for some  $a' < a^{NE}$  that is close to  $a^{NE}$ . The second part follows from the observation that any action  $a \in (a^{NE}, a^*]$  in the initial period can be replaced by  $a^{NE}$  without affecting any incentive constraints. Thus every worst punishment path starts with an action that is no more efficient than  $a^{NE}$ . But how does the path look like from the second period? Call any path  $(a^s, a^c, a^c, \dots)$  such that  $a^s < a^c$  a *stick-and-carrot path*. Abreu [1] shows that, with geometric discounting, the worst punishment payoff  $\underline{v}$  can be achieved by a stick-and-carrot punishment path with  $a^c = a^o$ , where one-period punishment is immediately followed by the best equilibrium path  $(a^o, a^o, \dots)$ . Furthermore it is the unique worst punishment path when  $\pi^*$  is increasing to the right of  $a^s$ .

We show that the worst punishment payoff can be achieved by a version of stick-and-carrot path where the play would eventually go back to the best equilibrium path when the players exhibit future bias in the MLRP sense. We also provide two conditions under which the worst punishment path is unique. The first condition is essentially the same as the above condition on the slope of  $\pi^*$ . The second condition, which is new, is that the

future bias is strict in some sense. The bottom line is that the structure of stick-and-carrot punishment is robust with respect to future bias.

**Proposition 3.** *Suppose that  $\{\delta(t)\}_{t=1}^{\infty}$  exhibits future bias in the MLRP sense. Then the following hold:*

1. *There exists a worst punishment path  $\underline{a}^{\infty}$  that takes the following form:  
 $\underline{a}^{\infty} = (\underbrace{0, \dots, 0}_{K \text{ times}}, a, a^o, a^o, \dots)$ , where 0 appears  $K$  times for some finite  $K \geq 0$  and  $a \in [0, a^o]$ . A worst punishment path in the above form is unique (i.e.  $K$  and  $a$  are uniquely determined), and the incentive constraint to play  $\underline{a}^1$  on the path  $\underline{a}^{\infty}$  is binding, that is,*

$$\begin{aligned} \delta(1)(\pi^*(a) - \pi(a)) &= \delta(2)(\pi(a^o) - \pi(a)) \quad \text{when } K = 0, \text{ and} \\ \delta(1)(-\underline{\pi}) &= \delta(K+1)(\pi(a) - \underline{\pi}) + \delta(K+2)(\pi(a^o) - \pi(a)) \quad \text{when } K \geq 1. \end{aligned}$$

2. *The above path is the unique worst punishment path if (i)  $K = 0$  and  $\pi^*$  is increasing to the right of  $\underline{a}^1$  or (ii)  $\{\delta(t)\}_{t=1}^{\infty}$  is strictly more future biased relative to  $\{\delta_2(t)\}_{t=1}^{\infty}$  in the MLRP sense.*

The worst punishment path may require multiple periods of “stick” to generate a harsh enough punishment. This is due to the assumption that the stage game payoff is bounded. Otherwise this path would be a usual stick-and-carrot path with one period of stick. For the standard geometric discounting, which satisfies the assumption for the proposition, there are many other worst punishment paths. For example, a simple stick-and-carrot path  $(0, a', a', \dots)$  would be another worst punishment paths, where  $a'$  satisfies  $V(0, a', a', \dots) = V(0, \dots, 0, a, a^o, a^o, \dots)$ . The second part of the proposition shows that such multiplicity of punishment paths is eliminated by strict future bias. This is because backloading payoffs does make a punishment strictly more harsh in this case.

The idea of the proof is simple. We can replace any worst punishment path  $\underline{a}^{\infty}$  with the (unique) path that takes the above form and has an equal initial value. The only binding incentive constraint is the one in the first period. Since the first period action of this new path is no more than the first period action of the original path and  $\pi^*$  is nondecreasing, this new path is indeed an ASPE path. By construction,  $V(0, \dots, 0, a, a^o, a^o, \dots) = V(\underline{a}^{\infty})$ . Then the assumption of future bias implies  $c_2(0, \dots, 0, a, a^o, a^o, \dots) \leq c_2(\underline{a}^{\infty})(= \underline{v})$ . Hence this new ASPE path achieves the worst punishment value  $\underline{v}$  from the viewpoint of “period-0” player, against whom this punishment would be used upon his deviation.

If a player is strictly more future biased than his period-2 self and  $\underline{a}^{\infty}$  does not take this version of stick-and-carrot form, then we would have  $c_2(0, \dots, 0, a, a^o, a^o, \dots) < \underline{v}$  at

the end of the above argument, which is a contradiction. Hence the worst punishment path must take this particular form in this case and clearly there exists the unique such path to generate  $\underline{v}$ .

The assumption of strict future bias in the MLRP sense is not weak. But suppose that  $K = 0$  and the worst punishment payoff can be achieved by a standard stick-and-carrot path (this is the case, for example, if the capacity constraint is large enough in the Cournot game). Then what we would need for the uniqueness result is just that  $\frac{\delta(t_2)}{\delta(1)} > \frac{\delta_2(t_2)}{\delta_2(1)}$  for all  $t_2 > 1$  (i.e.  $t_1 = 1$  is enough). For example this condition is satisfied for  $\beta$ - $\delta$  discounting with  $\beta > 1$ , although  $\{\delta(t)\}_{t=1}^\infty$  is not strictly more future biased relative to  $\{\delta_2(t)\}_{t=1}^\infty$  in this case.

Once we characterize the structure of worst punishment paths, we can derive the set of equilibrium payoffs. For the case of future bias, the worst equilibrium path coincides with the worst punishment path characterized above. This is because the same argument as above implies that the worst equilibrium path can take exactly the same form as the worst punishment path, thus the worst equilibrium payoff is given by  $V(\underline{a}^\infty)$ . This is a trivial fact with geometric discounting, but does not hold in general with general discounting. The following proposition also proves the convexity of the equilibrium payoff set.

**Proposition 4.** *Suppose that  $\{\delta(t)\}_{t=1}^\infty$  exhibits future bias in the MLRP sense. Let  $\underline{a}^\infty$  be the worst punishment path in the form described in Proposition 3. Then the set of equilibrium payoffs is given by the closed interval  $[\underline{v}_1, \bar{v}]$ , where  $\underline{v}_1 = V(\underline{a}^\infty)$ .*

We can show that any payoff between  $\underline{v}_1$  and  $\bar{v}$  can be supported by a similar, but more generous stick and carrot (with  $K' \leq K$ ). We do not use a public randomization device to convexify the equilibrium payoff set.

Based on Proposition 3, we can compute the value of the worst punishment payoff for  $\{\beta_t\}_t$ -weighted discounting that is future biased in the MLRP sense and more future biased than the geometric discounting. For this class of discounting, which includes the geometric discounting as a special case, we can also show that the worst punishment payoff in the limit (as  $\delta \rightarrow 1$ ) is strictly smaller than the stage game minmax payoff 0 if and only if the discounting is not geometric.

**Proposition 5.** *Suppose that the time preference is given by  $\{\beta_t\}_t$ -weighted discounting. If  $\{\beta_t\}_t$ -weighted discounting exhibits future bias and is more future biased than the geometric discounting with the same  $\delta$  both in the MLRP sense, then the following hold:*

1.  $\beta_t$  is nondecreasing with respect to  $t$  and converging to  $\lim_{t \rightarrow \infty} \beta_t < \infty$ .
2. Let  $\underline{a}^\infty = (\underbrace{0, \dots, 0}_{K \text{ times}}, a, a^o, a^o, \dots)$  be the worst punishment path in the form described in Proposition 3, where  $K \geq 0$  and  $a \in [0, a^o)$ . The worst punishment payoff  $\underline{v}$  is given

by the following value:

$$\pi(a^o) - \frac{\pi^*(a) - \pi(a)}{\sum_{t=1}^{\infty} \beta_t \delta^t}$$

when  $K = 0$ , and

$$\frac{\delta}{(1 - \delta)(\sum_{t=1}^{\infty} \beta_t \delta^t)} \left\{ \sum_{t=0}^{K-1} [(\beta_{t+1} - \beta_t) \delta^t] \underline{\pi} + (\beta_{K+1} - \beta_K) \delta^K \pi(a) + \sum_{t=K+1}^{\infty} [(\beta_{t+1} - \beta_t) \delta^t] \pi(a^o) \right\}$$

when  $K \geq 1$ .  $K$  is nondecreasing as  $\delta$  increases and goes to  $\infty$  as  $\delta \rightarrow 1$ .

3. As  $\delta \rightarrow 1$ ,  $\underline{v}$  and  $\underline{v}_1$  converge to the same limit

$$\frac{\lim_t \beta_t - 1}{\lim_t \beta_t} \underline{\pi},$$

which is strictly smaller than the symmetric minmax payoff 0 if and only if the  $\{\beta_t\}_t$ -weighted discounting is not the geometric discounting.

A few remarks are in order. First, the third result follows from the fact that  $\lim_t \beta_t > 1$  if and only if the discounting is not geometric. To see this, note that the two MLRP assumptions imply  $\frac{\beta_t}{\beta_{t-1}} \geq \frac{\beta_{t+1}}{\beta_t}$  and  $\beta_t \geq 1$  for all  $t = 1, 2, \dots$  respectively. This implies that  $\beta_t$  must be nondecreasing with diminishing increasing rates. Hence  $\lim_t \beta_t = 1$  if and only if  $\beta_t = 1$  for all  $t \geq 1$  (remember that  $\beta_0 = 1$ ) if and only if the discounting is the geometric discounting (because  $\beta_t$  is bounded above). Otherwise  $\lim_t \beta_t > 1$ . Second, more importantly, an actual impact of bias on the equilibrium payoff can be huge. Note that all these results are stated in terms of normalized payoffs. In terms of total payoffs, both the worst punishment payoff and the worst equilibrium payoff diverge to negative infinity as  $\delta$  goes to 1 whenever the discounting is not geometric in this class.

We should be careful about how to interpret large  $\delta$  here. There are usually two ways to interpret this parameter. One interpretation is that players are more patient. The other interpretation is that players are playing the game more frequently. The first interpretation is more appropriate here because the bias structure is fixed. One advantage of our framework is that we can also model the second effect formally. In Section 6, we examine the limit equilibrium payoff when the length of periods becomes shorter while keeping the intertemporal bias structure unchanged with respect to actual time.

## 5 Characterization of Equilibrium: Present Bias

Next we examine the structure of equilibrium paths and the set of equilibrium payoffs for the case of present bias. This is possibly a more interesting case, but more complicated. In this section, we introduce the following additional assumptions on the stage game to facilitate our analysis.

**Assumption 2.**  $\pi$  is strictly concave and  $\pi^*$  is convex.

One implication of this assumption is that the set of  $a$  for which the constant play of  $a$  is an ASPE path is a closed interval  $[\hat{a}, a^o]$ , where  $\hat{a}$  is the unique action below  $a^{NE}$  such that the incentive constraint is binding, i.e.  $\delta(1)(\pi^*(\hat{a}) - \pi(\hat{a})) = (1 - \delta(1))(\pi(\hat{a}) - \underline{v})$ . This follows from Assumption 2 because  $\pi(a) - \delta(1)\pi^*(a)$  is strictly concave.

We first focus on an interesting, but restricted class of equilibrium. Then we move on to a fully general case.

### 5.1 Forgiving Punishment and Forgiving Equilibrium

It makes sense to punish a deviator as soon as possible in order to make the punishment as effective as possible. The stick-and-carrot strategy takes this idea to the extreme: it punishes a deviator as much as possible immediately after his deviation, then forgives him afterward, i.e. goes back to the original cooperative path. In another word, any necessary punishment is as front loaded as possible.

We consider a particular class of punishment scheme: *forgiving punishment*, which is based on this idea, but more general than a simple stick-and-carrot scheme. It imposes the harshest punishment in the first period and then imposes less and less severe punishments over time. So a deviator is gradually forgiven along a forgiving punishment path. Formally, let  $A_f^\infty = \{a^\infty \in A^\infty : a^1 \leq a^2 \leq \dots\}$  be the set of paths where payoffs are nondecreasing over time. We say that a path is forgiving if it belongs to this set. In this subsection, we restrict attention to strongly symmetric equilibria that prescribe a path in  $A_f^\infty$  after every history. We call such equilibria *forgiving equilibria*. With geometric discounting, it is without loss of generality to focus on forgiving equilibria: every equilibrium payoff can be generated by a forgiving equilibrium. Let  $A_f^{ASPE}$  be the set of all forgiving equilibrium paths.

It is easy to show as in Proposition 1 that the best forgiving equilibrium path is given by a constant path, which is denoted by  $(a_f^o, a_f^o, \dots)$ . We define the worst forgiving punishment payoff by  $\underline{v}_f = \inf_{a^\infty \in A_f^{ASPE}} c_2(a^\infty)$ . Since  $(a^{NE}, a^{NE}, \dots) \in A_f^{ASPE}$ ,  $A_f^{ASPE}$  is nonempty. Since  $A_f^\infty$  is compact, we can show the existence of a forgiving equilibrium path that achieves  $\underline{v}_f$  as usual. We denote a worst forgiving punishment path by  $\underline{a}_f^\infty$ .

The following proposition characterizes the worst forgiving punishment path when the time preference exhibits (strict) present bias in the FOSD sense.

**Proposition 6.** *Suppose that  $\{\delta_2(t)\}_{t=1}^\infty$  is more present biased relative to  $\{\delta_3(t)\}_{t=1}^\infty$  in the FOSD sense.<sup>10</sup> Then the following holds:*

1. *There exists a worst forgiving punishment path that takes the following simple stick-and-carrot form:  $\underline{a}_f^\infty = (a^s, a^c, a^c, \dots)$  where  $a^s < a^c$ . Moreover,  $a^s$  is never the worst action 0 if  $\delta(1) > \delta_2(1)$  and  $\pi^{*'}(0)$  is small enough.*
2. *The above path is the unique worst forgiving punishment path if  $\{\delta_2(t)\}_{t=1}^\infty$  is strictly more present biased relative to  $\{\delta_3(t)\}_{t=1}^\infty$  in the FOSD sense.*

Take any forgiving punishment path  $a^\infty$ . Then we can find some  $a'$  to satisfy  $V(a^\infty) = V(a^1, a', a', \dots)$  and easily verify that  $(a^1, a', a', \dots)$  is another ASPE path. The present bias implies that  $(a^1, a', a', \dots)$  is at least as a harsh punishment as  $a^\infty$  (i.e.  $c_2(a^1, a', a', \dots) \leq c_2(a^\infty)$ ). Thus it is without loss of generality to focus on punishments of the form  $(a^s, a^c, a^c, \dots)$  for forgiving equilibria. This observation is valid with geometric discounting and not new.

However the way in which  $(a^s, a^c)$  is determined is very different from the standard case. With geometric discounting, the incentive constraint can be relaxed by making stick worse and carrot better while keeping the discounted average payoff constant, since  $\pi^*(a^s)$  decreases as  $a^s$  decreases. Hence the gap between stick and carrot is maximized, and therefore either 0 is used as stick or  $a^o$  is used as a carrot. With present bias, we also see the opposite force at work, which tends to reduce the gap between stick and carrot. If we increase  $a^s$  and decrease  $a^c$  in such a way that  $V(a^s, a^c, a^c, \dots)$  does not change, then this path generates a more harsh punishment if it is still an ASPE. Intuitively, due to present bias, we can decrease  $a^c$  by relatively a large amount by increasing  $a^s$  only slightly. From a deviator's viewpoint in one period before,  $a^c$  decreases too much thus  $c_2$  goes down. This effect pushes  $a^s$  up and pushes  $a^c$  down. The incentive cost associated with this perturbation depends on how quickly  $\pi^*(a^s)$  increases as  $a^s$  increases. Thus, as claimed in the latter part of Proposition 6-1, if the slope of  $\pi^*$  is small at 0, then this incentive cost is relatively small hence  $a^s$  must be strictly positive.

To be more explicit, the pair of stick and carrot  $(a^s, a^c)$  that constitutes a worst forgiving

---

<sup>10</sup>This is equivalent to  $\delta_2(1) \geq \delta_\tau(1)$  for all  $\tau \geq 3$ .



punishment path can be obtained by solving the following minimization problem:

$$\begin{aligned} & \min_{a^s, a^c \in A} \delta_2(1)\pi(a^s) + (1 - \delta_2(1))\pi(a^c) \\ \text{s.t. } & \frac{\delta(1)}{1 - \delta(1)}(\pi^*(a^s) - \pi(a^s)) \leq \delta_2(1)(\pi(a^c) - \pi(a^s)), \\ & \frac{\delta(1)}{1 - \delta(1)}(\pi^*(a^c) - \pi(a^c)) \leq \delta_2(1)(\pi(a^c) - \pi(a^s)). \end{aligned}$$

The first constraint corresponds to the one-shot deviation constraint for  $a^s$  and the second constraint corresponds to the one-shot deviation constraint for  $a^c$ . When  $\delta(1)$  is small, we can drop the second constraint because the obtained solution  $a^s$  and  $a^c$  automatically satisfies  $a^s < a^c < a_f^o$ .

In the following proposition, we show that the worst forgiving equilibrium payoff can also be achieved by a (different) stick and carrot path and that the set of forgiving equilibrium payoffs is convex.

**Proposition 7.** *Suppose that  $\{\delta_2(t)\}_{t=1}^\infty$  is more present biased relative to  $\{\delta_3(t)\}_{t=1}^\infty$  in the FOSD sense. Then the following hold:*

1. *There exists a worst forgiving equilibrium path that takes the following form:  $\underline{a}_{1,f}^\infty = (\hat{a}^s, \hat{a}^c, \hat{a}^c, \dots)$  where  $\hat{a}^s \leq \hat{a}^c$ . If  $\delta(1) \geq \delta_2(1)$ , then  $\hat{a}^s \leq a^s < a^c \leq \hat{a}^c$  for any worst forgiving punishment path  $\underline{a}_f^\infty = (a^s, a^c, a^c, \dots)$  in the stick-and-carrot form.<sup>11</sup>*
2. *The set of forgiving equilibrium payoffs is given by a closed interval  $[\underline{v}_{1,f}, \bar{v}_{1,f}]$ , where  $\underline{v}_{1,f} = V(\underline{a}_{1,f}^\infty)$  and  $\bar{v}_{1,f} = \pi(a_f^o)$ .*
3. *All the three payoffs  $\underline{v}_f$ ,  $\underline{v}_{1,f}$ ,  $\bar{v}_{1,f}$  and the associated equilibrium paths are completely determined by  $\delta(1)$  and  $\delta_2(1)$  only.  $\underline{v}_f(\delta(1), \delta_2(1))$  decreases as  $\delta(1)$  decreases (while keeping  $\delta_2(1)$  constant) or as  $\delta_2(1)$  increases (while keeping  $\delta(1)$  constant).  $\underline{v}_{1,f}(\delta(1), \delta_2(1))$  decreases as  $\delta_2(1)$  increases.  $\bar{v}_{1,f}(\delta(1), \delta_2(1))$  increases up to  $\bar{\pi}$  as  $\delta(1)$  decreases or as  $\delta_2(1)$  increases.*

Propositions 7-3 implies that it is without loss of generality to use the standard  $\beta$ - $\delta$  discounting provided that the above assumption of present bias is satisfied. In another word, we cannot identify  $\beta_t, t \geq 2$  from equilibrium behavior as long as we focus on forgiving equilibria (even if we can learn what the players would have played off the equilibrium path). For  $\beta$ - $\delta$  discounting,  $\delta(1) = \frac{1-\delta}{1-\delta+\beta\delta}$  and  $\delta_2(1) = 1 - \delta$ . So, for any pair of  $(\delta(1), \delta_2(1)) \in (0, 1)^2$ , we can find appropriate  $\beta > 0$  and  $\delta \in (0, 1)$  to generate them, namely  $\beta = \frac{1-\delta(1)}{\delta(1)} \frac{\delta_2(1)}{1-\delta_2(1)}$  and  $\delta = 1 - \delta_2(1)$ . In particular, we have  $\beta \leq 1$  when  $\delta(1) \geq \delta_2(1)$ .

<sup>11</sup>If  $\delta(1) \leq \delta_2(1)$ , then  $\hat{a}^s \geq a^s$  and  $\hat{a}^c \leq a^c$ .

Proposition 1 of Chade, Prokopovych and Smith [3] shows that the set of equilibrium payoffs expands weakly as  $\beta$  increases and  $\delta$  decreases while  $\delta(1)$  is fixed. Our result is consistent with this proposition as such change of parameters induces  $\delta_2(1)$  to increase. Chade, Prokopovych and Smith [3] also shows that the set of equilibrium payoffs is nonmonotonic in either  $\beta$  or  $\delta$ . Our result is consistent with this non-monotonicity result as well. As  $\beta$  increases,  $\delta(1)$  becomes smaller while  $\delta_2(1)$  is unaffected, and the effect of  $\delta(1)$  on  $\underline{v}_{1,f}$  is ambiguous. As  $\delta$  increases,  $\delta(1)$  and  $\delta_2(1)$  both decrease. Again their effects on  $\underline{v}_{1,f}$  and  $\bar{v}_{1,f}$  are ambiguous.

As before, we can characterize the worst forgiving punishment payoff given  $\{\beta_t\}_t$ -weighted discounting when  $\delta$  approaches 1.

Let  $a^{s*}$  be a solution of the following minimization problem:

$$\min_{a \in A} \pi(a) + C(\{\beta_t\}_t) (\pi^*(a) - \pi(a)),$$

where  $C(\{\beta_t\}_t) = \frac{1 - \frac{\beta_1}{\sum_{t=1}^{\infty} \beta_t}}{\beta_1}$ . Define  $a^{c*}$  by the following equation:

$$\pi(a^{c*}) = \pi(a^{s*}) + \frac{1}{\beta_1} (\pi^*(a^{s*}) - \pi(a^{s*})). \quad (1)$$

**Proposition 8.** *Suppose that the time preference is given by  $\{\beta_t\}_t$ -weighted discounting that exhibits present bias in the MLRP sense and is not the geometric discounting. Then the following hold:*

1.  $a^{s*}$  and  $a^{c*}$  are uniquely determined, and  $\pi(a^{s*}) + C(\{\beta_t\}_t) (\pi^*(a^{s*}) - \pi(a^{s*})) > 0$  and  $\pi(a^{c*}) > 0$ .
2. The worst forgiving punishment path that is obtained in Proposition 6 converges to  $(a^{s*}, a^{c*}, a^{c*}, \dots)$  as  $\delta \rightarrow 1$ .
3.  $\underline{v}_f$  converges to  $\pi(a^{s*}) + C(\{\beta_t\}_t) (\pi^*(a^{s*}) - \pi(a^{s*}))$  as  $\delta \rightarrow 1$ . If  $\sum_{t=1}^{\infty} \beta_t = \infty$ , then the limit coincides with  $\pi(a^{c*})$ .

When  $\delta$  is large,  $\delta(1)$  is small. Hence we only need to consider the one-shot deviation constraint for  $a^s$ , which can be written as  $\pi(a^c) \geq \pi(a^s) + \frac{1}{\beta_1 \delta} (\pi^*(a^s) - \pi(a^s))$ . This constraint must be binding. Hence we can use this to eliminate  $a^c$  from the objective function. As  $\delta \rightarrow 1$ , the objective function to be minimized (to find  $a^s$ ) converges to the objective function for the minimization problem for  $a^{s*}$ . Then the limit of the solutions for the minimization problems as  $\delta \rightarrow 1$  converges to the solution of the limit minimization problem, which is  $a^{s*}$ .  $a^{c*}$  is obtained by computing the limit of the binding constraint (1).

Proposition 8 implies the the worst punishment payoff for the class of forgiving equilibria is strictly positive (hence above the the minmax payoff 0) for any present-biased discounting that is not geometric discounting. The  $\{\beta_t\}_t$ -weighted discounting is present biased if and only if  $\frac{\beta_t}{\beta_{t-1}} \leq \frac{\beta_{t+1}}{\beta_t}$  for all  $t$ . It coincides with geometric discounting (with discount factor  $\beta\delta$ ) if and only if  $\frac{\beta_{t+1}}{\beta_t}$  is some constant  $\beta \in (0, 1]$ , since  $\{\beta_t\}$  must be bounded above by assumption. Then  $\{\beta_t\}_t$ -weighted discounting is not geometric discounting if and only if one of the above inequalities hold strictly. In this case, it can be shown that  $C(\{\beta_t\}_t) > 1$ , which implies that  $\underline{v}_f > 0$  even as  $\delta$  goes to 1.<sup>12</sup>

This does not mean that the true worst punishment payoff is strictly positive because  $\underline{v}$  is in general smaller than  $\underline{v}_f$  as we will see. However we will show that  $\underline{v} = \underline{v}_f > 0$  holds for  $\beta\text{-}\delta$  discounting with  $\beta \in (0, 1)$ , which is present biased.

## 5.2 General Case

Next we characterize the set of ASPE without any restriction. A question of interest is whether restricting attention to forgiving punishments is without loss of generality (i.e.  $\underline{v}_f = \underline{v}$ ) or not. The following proposition partially characterizes the structure of the worst punishment path with present bias and shows that a forgiving punishment path is never a worst punishment path.

**Proposition 9.** *Suppose that  $\{\delta(t)\}_{t=1}^\infty$  exhibits present bias in the FOSD sense and that  $\delta_t(1)$  is decreasing in  $t$ . Suppose also that  $\delta(1)$  is small enough. Let  $\underline{a}^\infty$  be any worst punishment path. If  $\underline{a}^{t,\infty}$  is a forgiving path for some  $t$ , then  $\underline{a}^{t,\infty}$  must be the worst constant ASPE path  $(\hat{a}, \hat{a}, \dots)$  and  $t$  must be at least 3.*

The assumption about  $\{\delta(t)\}_{t=1}^\infty$  is weaker than assuming that  $\{\delta(t)\}_{t=1}^\infty$  exhibits strict present bias in the FOSD sense, since it would imply that  $\delta_t(1)$  is decreasing in  $t$  (and more).

This proposition suggests that a worst punishment path may exhibit a cyclical pattern. Later in this subsection, we show an example with present bias where the unique worst punishment path must fluctuate forever generically, although  $\delta_t(1)$  is only weakly decreasing and (strictly) decreasing only for  $t = 1, 2, 3$ .

Next we develop a recursive method to characterize and derive worst punishment paths. For this purpose, we restrict our attention to  $\{\beta_t\}_t$ -weighted discounting where bias disappears after  $T$  periods (i.e.  $\beta_T = \beta_{T+k}$  for all  $k = 1, 2, \dots$ ). We assume that  $T \geq 2$  and treat the case of  $T = 1$  (i.e.  $\beta\text{-}\delta$  discounting) separately in the next subsection. Since we have geometric discounting from period  $T + 1$  on, we have  $c_{T+1} = c_{T+2} = \dots$  and represent this

<sup>12</sup>For non-geometric discounting,  $\beta_1$  must be less than 1. If  $\sum_{t=1}^\infty \beta_t = \infty$ , then  $C(\{\beta_t\}_t) = \frac{1}{\beta_1} > 1$ . If  $\sum_{t=1}^\infty \beta_t < \infty$ , we can show that  $1 - \frac{\beta_1}{\sum_{t=1}^\infty \beta_t} > \beta_1$ , hence  $C(\{\beta_t\}_t) > 1$  still holds.

common function by  $c$ , which is given by  $c(a^\infty) = (1 - \delta)[\pi(a^1) + \delta\pi(a^2) + \delta^2\pi(a^3) + \dots]$ . Thus the time preference has a recursive component at the end, which allows us to apply the recursive formulation in Obara and Park [8].<sup>13</sup> We work with stage game payoffs  $\pi$  instead of actions  $a$ . Let  $V^\dagger = [\underline{\pi}, \bar{\pi}]$  be the set of feasible payoffs in the stage game.

The central idea of Obara and Park [8] is to decompose a sequence of stage game payoffs in the future into a nonrecursive part and a recursive part. In this particular case, we decompose them into the stage game payoffs from the 2nd period to the  $T$ th period and the continuation payoff from the  $(T + 1)$ th period that is evaluated by the standard geometric discounting (which we call the *continuation score*). Consider all such combinations of  $T - 1$  payoffs from the 2nd period to the  $T$ th period and a continuation score from the  $(T + 1)$ th period that can arise in equilibrium. We can represent it as a correspondence as follows:

$$W^*(\pi^{1,T-1}) = \{c(a^{T,\infty}) : a^\infty \in A^{SSASPE}, \pi(a^{1,T-1}) = \pi^{1,T-1}\},$$

where we use  $\pi(a^{t,t+k})$  to denote  $(\pi(a^t), \dots, \pi(a^{t+k}))$ .

As in Obara and Park [8], we consider an operator  $B$  which maps a correspondence  $W : (V^\dagger)^{T-1} \rightrightarrows V^\dagger$  into another correspondence  $W' : (V^\dagger)^{T-1} \rightrightarrows V^\dagger$  as follows. Given the correspondence  $W$ , define

$$\begin{aligned} \underline{w} &= \min_{\pi^{1,T-1}, v} \delta_2(1)\pi^1 + \dots + \delta_2(T-1)\pi^{T-1} + [1 - \delta_2(1) - \dots - \delta_2(T-1)]v \\ \text{s.t. } &\pi^{1,T-1} \in [\underline{\pi}, \bar{\pi}]^{T-1}, v \in W(\pi^{1,T-1}). \end{aligned} \quad (2)$$

This is the harshest punishment according to  $W$ . Then the new correspondence  $W'$  is defined by

$$\begin{aligned} W'(\pi^{1,T-1}) &= \left\{ (1 - \delta)\pi + \delta v : \right. \\ &\quad \delta(1)\pi^1 + \dots + \delta(T-1)\pi^{T-1} + \delta(T)\pi + [1 - \delta(1) - \dots - \delta(T)]v \\ &\quad \left. \geq \delta(1)\pi^d(\pi^1) + (1 - \delta(1))\underline{w}, \underline{\pi} \leq \pi \leq \bar{\pi}, v \in W(\pi^{2,T-1}, \pi) \right\}, \end{aligned} \quad (3)$$

where  $W'(\pi^{1,T-1}) = \emptyset$  if there is no pair  $(\pi, v)$  satisfying all the constraints in (3).  $\pi^d(\pi)$  denotes  $\pi^*(a)$  at  $a$  such that  $\pi(a) = \pi$ .  $W'$  represents the set of combinations of payoffs from the 1st period to the  $(T - 1)$ th period and a continuation score from the  $T$ th period that can be supported when we can pick “continuation payoffs” from  $W$ . Obara and Park [8] shows that  $W^*$  is the largest fixed point of the operator  $B$  and that the graph of  $W^*$  (i.e.  $\{(\pi^{1,T-1}, v) : \pi^{1,T-1} \in [\underline{\pi}, \bar{\pi}]^{T-1}, v \in W^*(\pi^{1,T-1})\}$ ) is nonempty and compact. It

<sup>13</sup>Obara and Park [8] provides a recursive characterization of equilibrium payoffs for a more general class of time preferences which is *eventually recursive*.

is easy to see that  $a^\infty$  is an ASPE path if and only if the corresponding payoffs satisfy  $v^{t+T-1} \in W^*(\pi^{t,t+T-2})$  for every  $t = 1, 2, \dots$ , where  $v^{t+T-1} = c(a^{t+T-1}, \infty)$ . Hence we can characterize all the ASPE paths once  $W^*$  is obtained.

In the next proposition, we show that the operator  $B$  preserves convexity.<sup>14</sup>

**Proposition 10.** *Let  $W$  be a correspondence from  $(V^\dagger)^{T-1}$  to  $V^\dagger$ . If the graph of  $W$  is nonempty and convex, then so is the graph of  $W' = B(W)$ .*

Since convexity is preserved under the operator  $B$  and we can start from a convex set  $(V^\dagger)^T$  to obtain the graph of  $W^*$  in the limit, we have the following result.

**Corollary 1.** *The graph of  $W^*$  is convex.*

Corollary 1 has many useful implications. Since the graph of  $W^*$  is convex, using public randomization does not expand the set of equilibrium payoffs. Also, the set  $dom(W^*) := \{\pi^{1,T-1} \in [\underline{\pi}, \bar{\pi}]^{T-1} : W^*(\pi^{1,T-1}) \neq \emptyset\}$  is convex, and  $W^*$  can be described by two functions  $\underline{W}^*(\pi^{1,T-1})$  and  $\overline{W}^*(\pi^{1,T-1})$  defined on  $dom(W^*)$  representing the lower and upper boundaries of  $W^*$  respectively, so that  $W^*(\pi^{1,T-1}) = [\underline{W}^*(\pi^{1,T-1}), \overline{W}^*(\pi^{1,T-1})]$  for all  $\pi^{1,T-1} \in dom(W^*)$ . Since the graph of  $W^*$  is convex,  $\underline{W}^*(\pi^{1,T-1})$  is convex and  $\overline{W}^*(\pi^{1,T-1})$  is concave. In fact, the upper boundary  $\overline{W}^*(\pi^{1,T-1})$  is constant at  $\bar{v}$ , because if  $a^\infty$  is an ASPE then so is  $(a^{1,t}, a^o, a^o, \dots)$  for any  $t$ . Since  $\underline{W}^*(\pi^{1,T-1})$  is convex and the graph of  $W^*$  is compact,  $\underline{W}^*(\pi^{1,T-1})$  is continuous in each argument.

For the following main result in this section, we introduce a weak form of present bias to  $\{\beta_t\}_t$ -weighted discounting, namely  $\beta_{T-1} > \beta_T$  and show that the lower boundary must be almost self-generating. This result leads to a simple algorithm to derive a worst punishment path.

**Proposition 11.** *Assume  $\{\beta_t\}_t$ -weighted discounting such that  $\beta_{T-1} > \beta_T = \beta_{T+1} = \dots$  for some  $T \geq 2$ . Then the following properties hold.*

1. *For any  $v = \underline{W}^*(\pi^{1,T-1})$  for any  $\pi^{1,T-1}$ ,  $(\pi', v')$  that decomposes  $v$  (according to (3)) must satisfy either (i)  $v' = \underline{W}^*(\pi^{2,T-1}, \pi')$  or (ii)  $\pi' = \bar{\pi}$ .*
2. *If  $a^\infty$  is a worst punishment path, then  $(\pi^{1,T-1}, v^T)$  solves (2) with respect to  $W^*$  and  $v^{t+T-1} = \underline{W}^*(\pi^{t,t+T-2})$  must hold for every  $t = 1, 2, \dots$  provided that  $a^t < a^*$  for every  $t$ .*
3. *Take any  $(\pi^{1,T-1}, v^T)$  that solves (2) with respect to  $W^*$ . For each  $v^{t+T-1}, t = 1, 2, \dots$ , let  $(\pi^{t+T-1}, v^{t+T})$  be the pair of  $(\pi', v')$  that satisfies  $v^{t+T-1} = (1-\delta)\pi' + \delta v'$  and  $v' \in$*

---

<sup>14</sup>Note that  $\pi^d(\pi)$  is convex, since  $\pi(a)$  is concave and  $\pi^*(a)$  is convex&nondecreasing.

$W^*(\pi^{t+1, t+T-2}, \pi')$ , where  $\pi^{t+T-1}$  is the largest among such  $(\pi', v')$ . Then the corresponding path  $a^\infty$  is a worst punishment path, where either (i)  $v^{t+T-1} = \underline{W}^*(\pi^{t, t+T-2})$  or (ii)  $\pi^{t+T-2} = \bar{\pi}$  holds for every  $t = 1, 2, \dots$

The first result is the key observation. It says that every point on the lower boundary must be decomposed with respect to the lower boundary or possibly the right edge of  $W^*$ . The reason is that if  $(\pi^{2, T-1}, \pi', v')$  is not on the lower boundary and  $\pi'$  is not  $\bar{\pi}$ , then we can increase  $\pi'$  and decrease  $v'$  to obtain an equilibrium starting with payoffs  $\pi^{1, T-1}$  and a continuation score lower than  $v$ , which is a contradiction.<sup>15</sup> If the most efficient action cannot be supported (i.e.  $a^o < a^*$ ), then  $\pi' = \bar{\pi}$  cannot hold, hence the lower boundary must be self-generating in this case. The fact that the boundary of  $W^*$  is self-generating would be helpful for us to derive  $W^*$  analytically or numerically.

The second result follows from the first result almost immediately. The initial point to achieve the worst punishment payoff  $\underline{v}$  must be on the lower boundary. Since  $\pi^t$  is never  $\bar{\pi}$  by assumption, we can apply the first result repeatedly to show that all the subsequent points are also on the lower boundary.

Once we obtain the lower boundary  $\underline{W}^*$ , we have a simple algorithm to construct a worst punishment path  $\underline{\pi}^\infty$  as suggested by the last result. First, we solve (2) to obtain an initial point  $(\underline{\pi}^{1, T-1}, v^T)$  where  $v^T = \underline{W}^*(\underline{\pi}^{1, T-1})$ . Next, we decompose  $v^T$  into  $(\underline{\pi}^T, v^{T+1})$  as follows. We take the intersection of  $(1 - \delta)\pi + \delta v = v^T$  and  $v \in W^*(\underline{\pi}^{2, T-1}, \pi)$ , which is a closed interval and nonempty by definition. Then pick  $(\pi, v)$  where  $\pi$  is the largest within this set and define it as  $(\underline{\pi}^T, v^{T+1})$ . Since the incentive constraint in (3) is satisfied for some  $(\pi, v)$  in this set, it is guaranteed to be satisfied for  $(\underline{\pi}^T, v^{T+1})$ . Then we can continue to decompose  $v^{T+1}$  and obtain an entire path  $\underline{\pi}^\infty$  that satisfies  $v^{t+T-1} \in W^*(\pi^{t, t+T-2})$  for all  $t = 1, 2, \dots$  and all the one-shot deviation constraints. This path is an ASPE path, and clearly a worst punishment path.

Note that the decomposition at each step would be usually simpler than it appears. The above intersection of  $(1 - \delta)\pi + \delta v = v^T$  and  $v \in W^*(\underline{\pi}^{2, T-1}, \pi)$  is typically a point: it may not be a singleton only if the lower boundary  $\underline{W}^*(\underline{\pi}^{2, T-1}, \pi)$  has a flat part that is parallel to  $(1 - \delta)\pi + \delta v = v^T$ .<sup>16</sup>

To see how this works, we consider the special case where  $T = 2$ . With  $T = 2$ , the considered  $\{\beta_t\}_t$ -weighted discounting can be expressed as  $\beta - \gamma - \delta$  discounting where  $\beta_1 = \beta$  and  $\beta_2 = \beta_3 = \dots = \beta\gamma$ . If  $\gamma < 1$ , we can apply Proposition 11. With  $T = 2$ ,  $\underline{W}^*(\pi^{1, T-1})$  is a function of a single variable, and thus it is easier to visualize the decomposition of a worst punishment path, as illustrated in Figure 2. It can be shown that the lower boundary

<sup>15</sup>Note that  $\pi'$  is not  $\bar{v}$  when  $\pi' \neq \bar{\pi}$  and  $W^*(\pi^{2, T-1}, \pi')$  is not a singleton  $\{\bar{v}\}$ . If  $(\pi^{2, T-1}, \pi', v')$  is not on the lower boundary, then  $v' > \underline{W}^*(\pi^{2, T-1}, \pi')$ , hence  $W^*(\pi^{2, T-1}, \pi')$  is not a singleton. Therefore such perturbation is feasible given  $\pi' < \bar{\pi}$ .

<sup>16</sup>Uniqueness can be guaranteed when  $T = 2$  and  $\pi^d(\pi)$  is strictly convex on  $dom(W^*)$ .

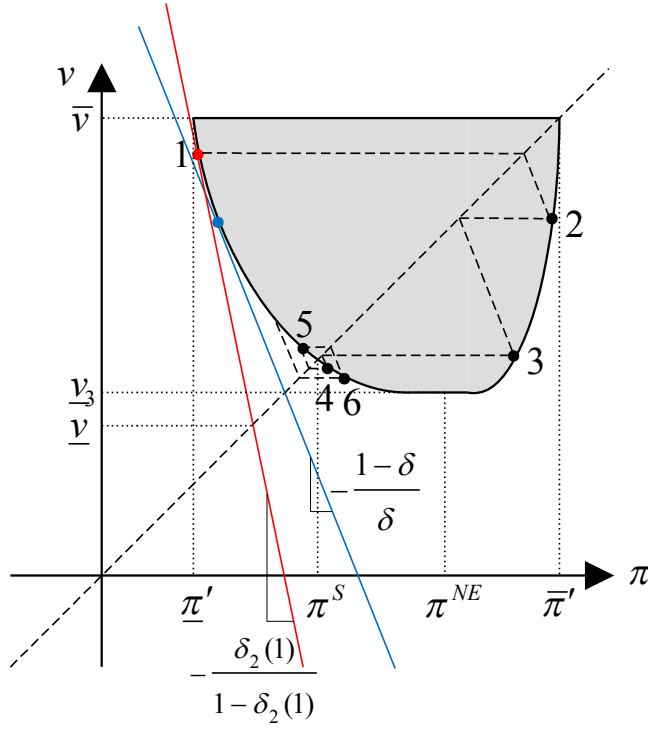


Figure 2: Illustration of the graph of  $W^*(\pi)$ .

is U-shaped with minimum attained at  $\pi(a^{NE}) = \pi^{NE}$ . Let  $(\pi^S, \pi^S)$  be the intersection of the lower boundary  $v = \underline{W}^*(\pi)$  and the 45-degree line  $v = \pi$ , where  $\pi^S = \pi(\hat{a})$  corresponds to the payoff at the worst constant equilibrium path.

The initial payoff  $\underline{\pi}^1$  in the worst punishment path  $\underline{\pi}^\infty$  can be obtained by solving  $\min_{\pi \in [\underline{\pi}', \bar{\pi}']} \delta_2(1)\pi + (1 - \delta_2(1))\underline{W}^*(\pi)$ , and the initial point  $(\underline{\pi}^1, \underline{W}^*(\underline{\pi}^1))$  is depicted as point 1 in Figure 2. The subsequent payoffs  $\underline{\pi}^2, \underline{\pi}^3, \dots$  (depicted as points 2, 3, and so on in Figure 2) cycle around  $\pi^S$ . Starting from a payoff  $\underline{\pi}^1$  lower than  $\pi^S$ , the payoff in the next period becomes higher than  $\pi^S$ , gradually declines below  $\pi^S$ , and then jumps back to a payoff higher than  $\pi^S$ , and this pattern repeats itself indefinitely. (In Figure 2, we have  $\underline{\pi}^2 > \underline{\pi}^3 > \underline{\pi}^4 > \pi^S > \underline{\pi}^5$  and  $\underline{\pi}^6 > \pi^S$ .) Thus the worst punishment path exhibits Edgeworth cycle-type price dynamics endogenously. The continuation path  $\underline{\pi}^{t,\infty}$  in the worst punishment path is monotone (in particular, a forgiving path) for some  $t$  only when we have  $\underline{\pi}^\infty = (\underline{\pi}', \bar{v}, \bar{v}, \dots)$  or we have  $\underline{\pi}^{t,\infty}$  decreasing initially, hit  $\pi^S$  exactly, and stay there forever. The latter case is a degenerate case but is shown to be possible in Proposition 9. The former case can be eliminated if  $\beta < \gamma$  (i.e. the case of present bias in the MRLP sense) and  $\delta$  is close to 1 by Proposition 9.

Finally, observe that this recursive method can be easily applied to the opposite case

with  $\beta_{T-1} < \beta_T = \beta_{T+1} = \dots$ . Proposition 10 and Corollary 1 still apply to this case. For this case with future bias, the worst punishment path is decomposed with respect to the upper boundary of  $W^*$  or possibly the left edge of  $W^*$ . The upper boundary of  $W^*$  is still a flat line at  $\bar{v}$ , hence this suggests that every worst punishment path must converge to the optimal path  $(a^o, a^o, \dots)$  eventually, which is consistent with our results for the case of future bias.

### 5.3 $\beta$ - $\delta$ Discounting

Consider quasi-hyperbolic discounting or  $\beta$ - $\delta$  discounting with  $\beta > 0$ . The next proposition shows that, with  $\beta$ - $\delta$  discounting, we can focus on stick-and-carrot paths without loss of generality when finding the worst punishment payoff.

**Proposition 12.** *Assume  $\beta$ - $\delta$  discounting where  $\beta > 0$  and  $\delta \in (0, 1)$ . Then there exists a worst punishment path  $\underline{a}^\infty$  of the form  $\underline{a}^\infty = (a^s, a^c, a^c, \dots)$  where  $a^s < a^c$ .*

The idea is similar to that behind forgiving punishments. Since the time preference becomes geometric discounting from the second period on, we can replace the path from the second period with a constant path without affecting the values of  $c_2$  and  $V$ . Hence the incentive to play the period-1 action  $a^1$  is preserved while the incentive constraint to play the constant path  $(a^c, a^c, \dots)$  is satisfied if  $\pi(a^1) \leq \pi(a^c)$ , which is the case in any worst punishment path. Thus the worst punishment payoff is achieved by a stick-and-carrot path. This result implies that, with  $\beta$ - $\delta$  discounting, it is without loss of generality to focus on forgiving punishments (i.e.  $\underline{v} = \underline{v}_f$ ).

When  $\beta > 1$ ,  $\beta$ - $\delta$  discounting exhibits future bias in the MLRP sense (but not strictly). We know that there is a worst punishment path of the form described in Proposition 3 in this case. By Proposition 5, the worst punishment payoff is  $\frac{\beta-1}{\beta}\underline{\pi} < 0$  when  $\delta$  is large. Proposition 12 shows that there is also another worst punishment path in the stick-and-carrot form  $(a^s, a^c, a^c, \dots)$ . In particular, if the worst punishment path obtained from Proposition 3 is not a stick-and-carrot path (which implies  $a^s = 0$ ), then we can construct a worst punishment  $(a^s, a^c, a^c, \dots)$ , where  $a^s = 0$  and  $a^c < a^o$  by smoothing out the payoffs from the second period.

When  $\beta < 1$ ,  $\beta$ - $\delta$  discounting exhibits present bias in the MLRP sense. Thus our results on forgiving punishments with present bias (Propositions 6–8) are applicable to this case. Proposition 6-1 suggests that the worst punishment path in the stick-and-carrot form may need to use a carrot  $a^c$  smaller than  $a^o$  even when the constraint  $a^s \geq 0$  is not binding. Also, by Proposition 8,  $\underline{v}(\delta)$  is strictly positive and bounded away from 0 as  $\delta \rightarrow 1$ .<sup>17</sup> From

---

<sup>17</sup>Note that this result holds under any other equilibrium concept that is more restrictive than ASPE, such as SPE.



Proposition 8, we can also see that  $\lim_{\delta \rightarrow 1} \underline{v}(\delta)$  is continuous and decreasing with respect to  $\beta$  and coincides with the minmax payoff 0 if and only if  $\beta = 1$ .

## 6 Characterization of Limit Equilibrium Payoff Sets

Consider the scenario where a repeated game begins at time 0 and the length of periods of the repeated game is given by  $\Delta > 0$ . The stage game in period  $t$  is played during the time interval  $[(t-1)\Delta, t\Delta]$  for  $t = 1, 2, \dots$ . We assume that the discount rate at time  $\tau \geq 0$  is given by

$$\rho(\tau) = r(1 - \eta(\tau))e^{-r(\tau - \int_0^\tau \eta(s)ds)},$$

where  $r > 0$  and  $\eta : \mathbb{R}_+ \rightarrow [0, 1)$  is a continuous function. In this way, we can introduce a fixed bias structure through  $\eta$  independent of the length of periods  $\Delta$ . So we can examine the effect of more frequent interactions while keeping the bias structure constant.

We assume that there exists some  $\tilde{T} > 0$  such that  $\eta(\tau)$  is decreasing on  $[0, \tilde{T}]$  and  $\eta(\tau) = 0$  for all  $\tau \geq \tilde{T}$ . If  $\tilde{T} = 0$  (hence  $\eta(t) = 0$  for all  $t$ ), then this model reduces to the standard exponential discounting. In the following, we assume that  $\eta(0) > 0$ , hence  $\tilde{T} > 0$ . As shown below, this is a continuous time version of future bias.

The discount factor for period  $t$  is given by

$$\begin{aligned} \delta(t) &= \int_{(t-1)\Delta}^{t\Delta} \rho(\tau) d\tau \\ &= e^{-r((t-1)\Delta - \int_0^{(t-1)\Delta} \eta(s)ds)} \left[ 1 - e^{-r(\Delta - \int_{(t-1)\Delta}^{t\Delta} \eta(s)ds)} \right] \end{aligned}$$

for all  $t = 1, 2, \dots$ . Since  $\rho(\tau) > 0$  for all  $\tau \geq 0$  and  $\int_0^\infty \rho(\tau) d\tau = 1$ , we have  $\delta(t) > 0$  for all  $t = 1, 2, \dots$  and  $\sum_{t=1}^\infty \delta(t) = 1$ . Note that

$$\frac{\delta(t+1)}{\delta(t)} = e^{-r(\Delta - \int_{(t-1)\Delta}^{t\Delta} \eta(s)ds)} \frac{1 - e^{-r(\Delta - \int_{t\Delta}^{(t+1)\Delta} \eta(s)ds)}}{1 - e^{-r(\Delta - \int_{(t-1)\Delta}^{t\Delta} \eta(s)ds)}}.$$

Using the properties of  $\eta$ , we can show that

$$\frac{\delta(2)}{\delta(1)} > \frac{\delta(3)}{\delta(2)} > \dots > \frac{\delta(T+1)}{\delta(T)} > \frac{\delta(T+2)}{\delta(T+1)} = \frac{\delta(T+3)}{\delta(T+2)} = \dots = e^{-r\Delta},$$

where  $T$  is the smallest integer such that  $T\Delta \geq \tilde{T}$ . So  $\{\delta(t)\}_{t=1}^\infty$  derived from  $\rho(\tau)$  exhibits future bias in the MLRP sense given any period length  $\Delta$ .

The following proposition characterizes the worst punishment payoff as  $\Delta \rightarrow 0$  in this case with future bias.

**Proposition 13.** *Assume the continuous time discounting  $\rho(\tau)$  with period length  $\Delta$ . Then as  $\Delta \rightarrow 0$ , the worst punishment payoff converges to*

$$\underline{v} \rightarrow \frac{\eta(0) - \eta(\widehat{T})}{1 - \eta(\widehat{T})} \underline{\pi} < 0,$$

where  $\widehat{T} > 0$  is the unique number satisfying

$$(1 - \eta(\widehat{T}))e^{-r(\widehat{T} - \int_0^{\widehat{T}} \eta(s)ds)} = (1 - \eta(0))\frac{-\underline{\pi}}{\bar{\pi} - \underline{\pi}}.$$

Since  $\eta(\widehat{T}) < \eta(0) < 1$ , we have  $\underline{\pi} < \lim_{\Delta \rightarrow 0} \underline{v} < 0$ . If  $\widehat{T}$  is large so that  $\widehat{T} \geq \widetilde{T}$ , then  $\eta(\widehat{T}) = 0$  and thus  $\underline{v} = \eta(0)\underline{\pi}$ . Note that  $\widehat{T}$  is affected by  $r$ , and when  $r$  is small,  $\widehat{T}$  would be large.

When  $\Delta$  is close to 0, the maximum profit  $\bar{\pi} = \pi(a^*)$  can be supported in equilibrium, which is clearly the best equilibrium payoff. In fact, we can support any payoff between the best equilibrium payoff and the worst punishment payoff as  $\Delta \rightarrow 0$  because any deviation would start a punishment instantly. Thus we have the following ‘‘folk theorem’’ as a corollary of Proposition 13.

**Corollary 2** (Folk Theorem). *Assume the continuous time discounting  $\rho(\tau)$  with period length  $\Delta$ . For any  $v \in \left(\frac{\eta(0) - \eta(\widehat{T})}{1 - \eta(\widehat{T})} \underline{\pi}, \bar{\pi}\right]$ , there exists  $\bar{\Delta} > 0$  such that  $v$  is an equilibrium (continuation) payoff for all  $\Delta \in (0, \bar{\Delta}]$ .*

It is useful to examine what would happen when we increase players’ patience (i.e. send  $r$  to 0) while fixing the period length  $\Delta$ . In this case, we obtain the following result.

**Proposition 14.** *Assume the continuous time discounting  $\rho(\tau)$  with period length  $\Delta$ . Then as  $r \rightarrow 0$ , the worst punishment payoff converges to*

$$\underline{v} \rightarrow \frac{1}{\Delta} \left( \int_0^{\Delta} \eta(s)ds \right) \underline{\pi} < 0.$$

Note that  $\widetilde{T}$  does not matter in this case because the timing of switch from  $\underline{\pi}$  to  $\bar{\pi}$  goes to  $\infty$  as the players become more and more patient.

Thus we obtain two different worst punishment payoffs depending on whether we let  $\Delta \rightarrow 0$  or  $r \rightarrow 0$ . The marginal future bias at  $t = 0$  matters when  $\Delta \rightarrow 0$ , whereas the average future bias for the first period matters when  $r \rightarrow 0$ . Therefore the worst punishment payoff is lower in the former case. Intuitively, this is because what matters is the bias in the very first period of the game in both cases.

We can express present bias by using  $\rho(\tau) = r(1 + \eta(\tau))e^{-r(\tau + \int_0^{\tau} \eta(s)ds)}$ . However, with present bias in general, it is difficult to express the worst punishment path explicitly in a

simple form, as argued and illustrated before. If we restrict attention to forgiving equilibria, we can do a similar analysis by using Proposition 8. When  $\Delta$  or  $r$  is sufficiently small, stick in the worst forgiving punishment path can be obtained by solving

$$\min_{a^s \in [a^*, \bar{a}]} \pi(a^s) + \frac{1 - \delta(1) - \delta(2)}{1 - \delta(1)} \frac{\delta(1)}{\delta(2)} (\pi^*(a^s) - \pi(a^s)),$$

while carrot can be obtained from

$$\pi(a^c) = \pi(a^s) + \frac{\delta(1)}{\delta(2)} (\pi^*(a^s) - \pi(a^s)).$$

Note that  $\frac{1 - \delta(1) - \delta(2)}{1 - \delta(1)} \frac{\delta(1)}{\delta(2)} > 1$ , and it converges to 1 and  $\frac{1 + \frac{1}{\Delta} \int_0^\Delta \eta(s) ds}{1 + \frac{1}{\Delta} \int_\Delta^{2\Delta} \eta(s) ds}$  as  $\Delta \rightarrow 0$  and  $r \rightarrow 0$ , respectively. In particular, as  $\Delta \rightarrow 0$ ,  $a^s$  approaches  $\tilde{a}$ , where  $\tilde{a}$  is the largest action such that  $\pi^*(\tilde{a}) = 0$  from above,  $a^c$  approaches the action with zero payoff from above, and  $\underline{v}_f$  converges to 0. In this case, relative present bias for the first two periods matters.

## 7 Appendix: Proofs

### Proof of Proposition 1

*Proof.* Fix  $\tau$  and choose  $\bar{a}_\tau^\infty$ . Define  $a_\tau^o = \sup\{\bar{a}_\tau^1, \bar{a}_\tau^2, \dots\}$ . Then there exists a sequence  $\{\bar{a}_\tau^{t_k}\}_{k=1}^\infty$  that converges to  $a_\tau^o$ . Since  $(\bar{a}_\tau^{t_k}, \bar{a}_\tau^{t_k+1}, \dots)$  is an equilibrium path, for each  $k$  we have

$$\delta(1)\pi(\bar{a}_\tau^{t_k}) + (1 - \delta(1))c_2(\bar{a}_\tau^{t_k+1, \infty}) \geq \delta(1)\pi^*(\bar{a}_\tau^{t_k}) + (1 - \delta(1))\underline{v}.$$

Since  $\pi$  is increasing, we have  $\pi(a_\tau^o) \geq c_2(\bar{a}_\tau^{t_k+1, \infty})$  and thus

$$\delta(1)\pi(\bar{a}_\tau^{t_k}) + (1 - \delta(1))\pi(a_\tau^o) \geq \delta(1)\pi^*(\bar{a}_\tau^{t_k}) + (1 - \delta(1))\underline{v}.$$

Taking limits of both sides, we obtain

$$\delta(1)\pi(a_\tau^o) + (1 - \delta(1))\pi(a_\tau^o) \geq \delta(1)\pi^*(a_\tau^o) + (1 - \delta(1))\underline{v}.$$

This inequality implies that the constant path  $(a_\tau^o, a_\tau^o, \dots)$  is an equilibrium path. Thus,  $\bar{v}_\tau \geq \pi(a_\tau^o)$ . Suppose that  $\bar{a}_\tau^t < a_\tau^o$  for some  $t$ . Since  $\delta_\tau(t) > 0$  for all  $t$ , we have  $\bar{v}_\tau = c_\tau(\bar{a}_\tau^\infty) < \pi(a_\tau^o)$ , which is a contradiction. Hence,  $\bar{a}_\tau^\infty$  must be the constant path  $(a_\tau^o, a_\tau^o, \dots)$ .

$a_\tau^o$  solves  $\max_{a \in A} \pi(a)$  subject to  $\delta(1)(\pi^*(a) - \pi(a)) \leq (1 - \delta(1))(\pi(a) - \underline{v})$ . Since the constraint set is nonempty and closed and  $\pi$  is increasing, there exists a unique solution. Note that this optimization problem does not depend on  $\tau$ , and thus we can write the solution  $a_\tau^o$  as  $a^o$ , an action independent of  $\tau$ .

Since  $\underline{v} \leq \pi(a^{NE})$ , the incentive constraint  $\delta(1)(\pi^*(a) - \pi(a)) \leq (1 - \delta(1))(\pi(a) - \underline{v})$  is satisfied at  $a = a^{NE}$ . The derivative of the left-hand side at  $a = a^{NE}$  is zero, while that of the right-hand side is positive. Thus there exists  $\hat{a} > a^{NE}$  satisfying the constraint, which implies  $a^o > a^{NE}$ .  $a^*$  satisfies the constraint for sufficiently small  $\delta(1)$ , and thus we have  $a^o = a^*$  when  $\delta(1)$  is small enough.  $\square$

### Proof of Proposition 2

*Proof.* Consider the strategy where players start the game by playing a path  $(a', a^{NE}, a^{NE}, \dots)$ , where  $a' < a^{NE}$ , and restart the path whenever a unilateral deviation occurs. Such a strategy is an equilibrium if the following incentive constraint is satisfied:

$$\delta(1)\pi(a') + (1 - \delta(1))\pi(a^{NE}) \geq \delta(1)\pi^*(a') + (1 - \delta(1))v_2,$$

where  $v_2 = \delta_2(1)\pi(a') + (1 - \delta_2(1))\pi(a^{NE})$ . Since the derivative of  $[\pi^*(a) - \pi(a)]$  at  $a = a^{NE}$  is zero and that of  $[\pi(a^{NE}) - \pi(a)]$  at  $a = a^{NE}$  is negative, we can choose  $a' < a^{NE}$  such that

$$\delta(1)[\pi^*(a') - \pi(a')] < (1 - \delta(1))\delta_2(1)[\pi(a^{NE}) - \pi(a')].$$

Clearly this  $a'$  satisfies the above incentive constraint strictly. Hence  $(a', a^{NE}, a^{NE}, \dots)$  is an equilibrium path. Since  $a' < a^{NE}$ , we have  $c_\tau(a', a^{NE}, a^{NE}, \dots) < \pi(a^{NE})$  for all  $\tau$ .

Let  $\underline{a}_\tau^\infty$  be an equilibrium path that achieves  $\underline{v}_\tau$ . Suppose that  $\underline{a}_\tau^1 > a^{NE}$ . Then  $(a^{NE}, \underline{a}_\tau^{2,\infty})$  is an equilibrium path and generates a lower payoff than  $\underline{a}_\tau^\infty$ . This is a contradiction.  $\square$

### Proof of Proposition 3

*Proof.* We first show that there exists a worst punishment path in the form described in the proposition. Take any worst punishment path  $a^\infty$ . Since  $(0, 0, \dots)$  cannot be an equilibrium (continuation) path, we can choose  $t \geq 2$  such that  $a^t > 0$ . Suppose that there are infinitely many  $t' > t$  such that  $a^{t'} < a^o$ . We construct a new path  $\hat{a}^\infty$  as follows. Replace  $a^t$  with some  $\hat{a}^t \in (0, a^t)$  and  $a^{\tilde{t}}$  with  $a^o$  for every  $\tilde{t} \geq T$  such that  $a^{\tilde{t}} < a^o$  for some  $T > t$ . We can choose  $T$  and  $\hat{a}^t$  so that  $V(a^\infty) = V(\hat{a}^\infty)$ . By construction,  $\delta(t)(\pi(\hat{a}^t) - \pi(a^t)) + \sum_{\tilde{t} \geq T} \delta(\tilde{t})(\pi(a^o) - \pi(a^{\tilde{t}})) = 0$ . Take any integer  $\tau \in [2, t]$ . Note that  $\frac{\delta(\tilde{t})}{\delta(t)} = \frac{\delta_\tau(\tilde{t} - \tau + 1)}{\delta_\tau(t - \tau + 1)}$ , which is smaller than or equal to  $\frac{\delta(\tilde{t} - \tau + 1)}{\delta(t - \tau + 1)}$  by MLRP. Combined with the above equality, this implies  $\delta(t - \tau + 1)(\pi(\hat{a}^t) - \pi(a^t)) + \sum_{\tilde{t} \geq T} \delta(\tilde{t} - \tau + 1)(\pi(a^o) - \pi(a^{\tilde{t}})) \geq 0$ . Therefore we have  $V(\hat{a}^{\tau,\infty}) \geq V(a^{\tau,\infty})$  for any  $\tau = 2, \dots, t$ .

The incentive constraint at  $\hat{a}^t$  is satisfied because  $\pi^*$  is nondecreasing and the continuation payoff from period  $t$  is the same or larger. The incentive constraint for  $\hat{a}^\infty$  in any other period before period  $T$  is satisfied because the same action is played and the continuation

payoff is not worse. The incentive constraint from period  $T$  is satisfied because  $(a^o, a^o, \dots)$  is an ASPE path. Similarly we can show that  $c_2(\widehat{a}^\infty) \leq c_2(a^\infty)$ . Thus  $\widehat{a}^\infty$  is at least as harsh a punishment path as  $a^\infty$ , hence must be a worst punishment path.

So we can assume without loss of generality that the worst punishment path  $a^\infty$  converges to the best ASPE path  $(a^o, a^o, \dots)$  after finite periods. Suppose that  $a^\infty$  does not take the form described in the proposition. Then there must exist  $t$  and  $t' > t$  such that  $a^t > 0$  and  $a^{t'} < a^o$ . Take the smallest such  $t$  and the largest such  $t'$ . Decrease  $a^t$  and increase  $a^{t'}$  so that the present value does not change as before, until either  $a^t$  hits 0 or  $a^{t'}$  hits  $a^o$ . It can be shown exactly as before that the obtained new path is still a worst punishment path. Repeat this process with the new path. Since there are only finitely many such  $t$  and  $t'$ , we can eventually obtain a worst punishment path  $\underline{a}^\infty$  that takes the required form within a finite number of steps.

Next, we show that the incentive constraint to play  $\underline{a}^1$  on the path  $\underline{a}^\infty$  must be binding. Note that, since  $a^o > a^{NE}$  and  $\pi^*(a) - \pi(a)$  is nondecreasing when  $a > a^{NE}$ ,  $(a, a^o, a^o, \dots)$  is an equilibrium path for all  $a \in [a^{NE}, a^o]$ . Suppose that the incentive constraint in the first period is not binding. First consider the case of  $K = 0$ . We can reduce  $\underline{a}^2 = a^o$  and obtain a worse punishment payoff, a contradiction. Next consider the case of  $K \geq 1$ . If the incentive constraint in period 1 is not binding, then the incentive constraint in period  $t \leq K$  is clearly not binding with larger continuation payoffs. The incentive constraint in period  $K + 1$  is not binding either, due to the assumption that the deviation gain becomes larger as the action profile moves away from the Nash equilibrium. It follows from the incentive constraint in period  $t \leq K$  if  $\underline{a}^{K+1} \leq a^{NE}$  and from the incentive constraint for the best equilibrium path if  $\underline{a}^{K+1} \in (a^{NE}, a^o)$ . Then we can reduce  $\underline{a}^{K+2} = a^o$  to obtain a worse punishment payoff, a contradiction.

Now we show that the obtained path  $\underline{a}^\infty$  is the unique worst punishment path that takes this form. Note that  $c_2(\underbrace{0, \dots, 0}_{K \text{ times}}, a', a^o, a^o, \dots) > c_2(\underbrace{0, \dots, 0}_{K+1 \text{ times}}, a'', a^o, a^o, \dots)$  for any  $a', a'' \in [0, a^o)$  and  $K$  and that these values are increasing in  $a'$  and  $a''$  respectively. So there exists unique  $K \geq 0$  and  $a \in [0, a^o)$  such that  $c_2(a^\infty) = c_2(\underbrace{0, \dots, 0}_{K \text{ times}}, a, a^o, a^o, \dots)$ .

Lastly we show that the above path must be the unique worst punishment path given (i) or (ii). Suppose that there exists a worst punishment path  $a^\infty$  other than  $\underline{a}^\infty$ . Suppose that (i) is satisfied. Since  $K = 0$ ,  $\pi(a^t) \leq \pi(a^o)$  for all  $t \geq 2$  with strict inequality for some  $t$  and thus  $\pi(\underline{a}^1) < \pi(a^1)$ . Then  $\pi^*(\underline{a}^1) < \pi^*(a^1)$ , and the incentive constraint is not binding at  $\underline{a}^1$ , which is a contradiction. If (ii) is satisfied, then the strict bias implies that we have  $c_2(\underline{a}^\infty) < c_2(a^\infty)$  in the end of the above construction. Again this is a contradiction.  $\square$

#### Proof of Proposition 4

*Proof.* We can use the same proof as the proof of Proposition 3 to show that there exists a worst equilibrium path  $\underline{a}_1^\infty$  that takes exactly the same form as the worst punishment path  $\underline{a}^\infty$  described in Proposition 3. If  $\underline{a}_1^\infty$  is different from  $\underline{a}^\infty$ , then  $V(\underline{a}_1^\infty) < V(\underline{a}^\infty)$ . Then either  $\underline{a}_1^\infty$  has a larger  $K$  than  $\underline{a}^\infty$  has, or they have the same  $K$  but  $\underline{a}_1^\infty$  has a smaller  $(K+1)$ th component. But in either case  $c_2(\underline{a}_1^\infty) < c_2(\underline{a}^\infty)$ , which is a contradiction. Hence  $\underline{a}^\infty$  is a worst equilibrium path.<sup>18</sup>

Let  $\underline{a}^\infty = (0, \dots, 0, \underline{a}^{K+1}, a^o, a^o, \dots)$  be a worst punishment path. Note that the binding incentive constraint is the one in the first period. This implies that this path is an ASPE path even if we replace  $\underline{a}^{K+1}$  with any  $a \in [\underline{a}^{K+1}, a^o]$ . It is also easy to see that any path of the form:  $(\underbrace{0, \dots, 0}_{K' \text{ times}}, a, a^o, a^o, \dots)$  for any  $0 \leq K' < K$  and  $a \in A$  is an ASPE path. Hence we can generate any payoff in  $[V(\underline{a}^\infty), \pi(a^o)] = [\underline{v}_1, \bar{v}]$  by an ASPE path of this stick-and-carrot form.  $\square$

### Proof of Proposition 5

*Proof.* MLRP implies that  $\frac{\beta_1}{1} \geq \frac{\beta_2}{\beta_1} \geq \dots$ . If any of these ratios is strictly less than one, then  $\beta_t$  converges to 0. This contradicts the assumption that  $\{\beta_t\}_t$ -weighted discounting is more future biased than the geometric discounting with  $\delta$ . Hence all these ratios are at least as large as 1. This implies that  $1 \leq \beta_1 \leq \beta_2 \leq \dots$  and  $\beta_t$  is monotonically nondecreasing with respect to  $t$  and converging to some  $\lim_t \beta_t \geq 1$ .

Now we compute the value of the worst punishment path. Take a worst punishment path  $\underline{a}^\infty = (0, \dots, 0, a, a^o, a^o, \dots)$  in the form described in Proposition 3. First consider the case of  $K = 0$ , which corresponds to the case with small  $\delta$ . The incentive constraint at  $\underline{a}^1 = a$  is binding and given by

$$\pi^*(a) - \pi(a) = \beta_1 \delta (\pi(a^o) - \pi(a)). \quad (4)$$

Hence

$$\begin{aligned} c_2(\underline{a}^\infty) &= \pi(a^o) - \frac{\beta_1 \delta}{\sum_{t=1}^{\infty} \beta_t \delta^t} (\pi(a^o) - \pi(a)) \\ &= \pi(a^o) - \frac{\pi^*(a) - \pi(a)}{\sum_{t=1}^{\infty} \beta_t \delta^t}. \end{aligned}$$

Next consider the case of  $K \geq 1$ . Again the incentive constraint in period 1 is binding, and it can be represented as follows:

$$\pi^*(0) - \underline{\pi} = \beta_K \delta^K (\pi(a) - \underline{\pi}) + \beta_{K+1} \delta^{K+1} (\pi(a^o) - \pi(a)). \quad (5)$$

---

<sup>18</sup>The worst equilibrium path is not usually unique.

This shows that  $K$  needs to become arbitrary large as  $\delta$  goes to 1 because  $\beta_t \geq 1$  for all  $t$ . We can also write the incentive constraint in period 1 as follows:

$$V(\underline{a}^\infty) = (1 - \delta(1))c_2(\underline{a}^\infty),$$

where  $\delta(1) = \frac{1}{\sum_{t=0}^{\infty} \beta_t \delta^t}$  is the normalization factor. On the other hand, we can write  $V(\underline{a}^\infty)$  as follows:

$$\begin{aligned} V(\underline{a}^\infty) &= \delta(1)\underline{\pi} + (1 - \delta(1))c_2(\underline{a}^{2,\infty}) \\ &= \delta(1)\underline{\pi} + \frac{1 - \delta(1)}{\delta} \{c_2(\underline{a}^\infty) - (c_2(\underline{a}^\infty) - \delta c_2(\underline{a}^{2,\infty}))\} \\ &= \delta(1)\underline{\pi} + \frac{1 - \delta(1)}{\delta} \left\{ c_2(\underline{a}^\infty) - \delta_2(1)\underline{\pi} - \sum_{t=1}^{K-1} (\delta_2(t+1) - \delta_2(t)\delta) \underline{\pi} \right. \\ &\quad \left. - (\delta_2(K+1) - \delta_2(K)\delta) \pi(a) - \sum_{t=K+1}^{\infty} (\delta_2(t+1) - \delta_2(t)\delta) \pi(a^o) \right\}. \end{aligned}$$

Using  $\underline{v} = c_2(\underline{a}^\infty)$  and  $\delta_2(t) = \frac{\delta(1)}{1 - \delta(1)} \beta_t \delta^t$  for all  $t$ , we can solve these equations for  $\underline{v}$  as follows:

$$\begin{aligned} \underline{v} &= \frac{\delta}{1 - \delta} \frac{\delta(1)}{(1 - \delta(1))} \left\{ \sum_{t=0}^{K-1} [(\beta_{t+1} - \beta_t) \delta^t] \underline{\pi} + (\beta_{K+1} - \beta_K) \delta^K \pi(a) \right. \\ &\quad \left. + \sum_{t=K+1}^{\infty} [(\beta_{t+1} - \beta_t) \delta^t] \pi(a^o) \right\} \\ &= \frac{\delta}{(1 - \delta)(\sum_{t=1}^{\infty} \beta_t \delta^t)} \left\{ \sum_{t=0}^{K-1} [(\beta_{t+1} - \beta_t) \delta^t] \underline{\pi} + (\beta_{K+1} - \beta_K) \delta^K \pi(a) \right. \\ &\quad \left. + \sum_{t=K+1}^{\infty} [(\beta_{t+1} - \beta_t) \delta^t] \pi(a^o) \right\}. \end{aligned}$$

As  $\delta$  increases,  $\delta(1)$  decreases, and thus  $a^o$  increases until it becomes  $a^*$ . Then the incentive constraints shown in (4) and (5) become slack when  $\delta$  becomes larger. This means that we can support a more harsh punishment, which will reduce  $a$  first and then increase  $K$  when  $a$  becomes 0. Hence  $K$  is nondecreasing as  $\delta$  increases.

Next we evaluate the limit of this expression as  $\delta \rightarrow 1$ . Clearly the second term in the brackets converges to 0 as  $\delta \rightarrow 1$ . As for the third term, note that  $\sum_{t=K+1}^{\infty} [(\beta_{t+1} - \beta_t) \delta^t] \leq \sum_{t=\bar{K}}^{\infty} [(\beta_{t+1} - \beta_t)] = \lim_{t \rightarrow \infty} \beta_t - \beta_{\bar{K}}$  for any  $\bar{K}$  for any large enough  $\delta$ . Since the right-hand side converges to 0 as  $\bar{K} \rightarrow \infty$ , the third term must converge to 0 as well. For the same reason, we see that  $\lim_{\delta \rightarrow 1} \sum_{t=0}^{K-1} [(\beta_{t+1} - \beta_t) \delta^t] = \lim_{\delta \rightarrow 1} \sum_{t=0}^{\infty} [(\beta_{t+1} - \beta_t) \delta^t]$ .

Since we can reverse the order of limits on the right-hand side, the right-hand side is  $\sum_{t=0}^{\infty} (\beta_{t+1} - \beta_t) = \lim_t \beta_t - 1$ . Note also that  $(1 - \delta)(\sum_{t=1}^{\infty} \beta_t \delta^t) \rightarrow \lim_t \beta_t$ . Hence, as  $\delta \rightarrow 1$ , the above expression converges to  $\frac{\lim_t \beta_t - 1}{\lim_t \beta_t} \underline{\pi}$ . Since  $\underline{v}_1 = V(\underline{a}^\infty) = (1 - \delta(1))c_2(\underline{a}^\infty)$  and  $\delta(1) \rightarrow 0$ ,  $\underline{v}_1$  converges to the same limit as the limit of  $c_2(\underline{a}^\infty)$  as  $\delta \rightarrow 1$ . Since  $\lim_t \beta_t \geq 1$ , we have  $\frac{\lim_t \beta_t - 1}{\lim_t \beta_t} \underline{\pi} \leq 0$ . If the discounting is geometric, then  $\beta_t = 1$  for all  $t$  and thus  $\frac{\lim_t \beta_t - 1}{\lim_t \beta_t} \underline{\pi} = 0$ . If the discounting is not geometric, then  $\lim_t \beta_t > 1$  and thus  $\frac{\lim_t \beta_t - 1}{\lim_t \beta_t} \underline{\pi} < 0$ .  $\square$

### Proof of Proposition 6

*Proof.* Let  $\underline{a}_f^\infty$  be a worst forgiving punishment path. Since  $\underline{a}_f^\infty$  is a forgiving equilibrium path, we have  $\pi(\underline{a}_f^1) \leq c_2(\underline{a}_f^{2,\infty}) \leq \pi(a_f^o)$ . Define  $a^c = \pi^{-1}(c_2(\underline{a}_f^{2,\infty}))$ . Note that  $a^c \in [\underline{a}_f^1, a_f^o]$ . Define  $a^s = \underline{a}_f^1$ . We show that  $(a^s, a^c, a^c, \dots)$  is a forgiving equilibrium path. The incentive constraint to play  $a^s$  is satisfied by construction. We can check the incentive to play  $a^c$  on the path considering two cases. Suppose  $a^c \geq a^{NE}$ . By Assumption 2, the constant path  $(a^c, a^c, \dots)$  is supported by the punishment payoff  $\underline{v}_f$ . Suppose  $a^c < a^{NE}$ . Since  $a^c \geq a^s$ , we have  $\pi^*(a^c) - \pi(a^c) \leq \pi^*(a^s) - \pi(a^s)$ , and the incentive to play  $a^c$  on the path  $(a^c, a^c, \dots)$  is implied by the incentive to play  $a^s$  on the path  $(a^s, a^c, a^c, \dots)$ .

Since  $\{\delta_2(t)\}_{t=1}^\infty$  is more present biased relative to  $\{\delta_3(t)\}_{t=1}^\infty$  in the FOSD sense and the path  $\underline{a}_f^{2,\infty}$  is forgiving,  $c_2(\underline{a}_f^{2,\infty}) = \pi(a^c)$  implies  $c_3(\underline{a}_f^{2,\infty}) \geq \pi(a^c)$ . Hence  $c_2(a^s, a^c, a^c, \dots) \leq c_2(\underline{a}_f^\infty) = \underline{v}_f$ . Thus  $(a^s, a^c, a^c, \dots)$  is a worst forgiving punishment path.

Next we show that  $a^s < a^c$ . Since the equilibrium constructed in the proof of Proposition 2 is a forgiving equilibrium, we have  $\underline{v}_f < \pi(a^{NE})$ . Suppose  $a^s = a^c$ . Then the incentive to play  $a^s$  implies  $a^s = a^c = a^{NE}$ . This means  $\underline{v}_f = \pi(a^{NE})$ , a contradiction.

Suppose that  $\delta(1) > \delta_2(1)$  and  $a^s = 0$  for the stick-and-carrot path. If we increase  $a^s$  and decrease  $a^c$  slightly so that  $V(a^s, a^c, a^c, \dots)$  does not change, then  $\pi(\Delta a^s)\delta(1) + (1 - \delta(1))\pi(\Delta a^c) = 0$  is satisfied for  $\Delta a^s > 0$  and  $\Delta a^c < 0$ . Since  $\delta(1) > \delta_2(1)$ , this implies that  $\pi(\Delta a^s)\delta_2(1) + (1 - \delta_2(1))\pi(\Delta a^c) = \Delta \underline{v}_f < 0$ , that is, this perturbation generates a harsher punishment if the obtained path is still an ASPE path. We show that this path is indeed an ASPE for a small perturbation if  $\pi^{*'}(0) \geq 0$  is small. The same argument as above shows that  $(a^c + \Delta a^c, a^c + \Delta a^c, \dots)$  is an ASPE path for a small  $\Delta a^c < 0$ . The original incentive constraint for  $a^s = 0$  is  $V(0, a^c, a^c, \dots) \geq \delta(1)\pi^*(0) + (1 - \delta(1))\underline{v}_f$ . As we slightly increase  $a^s$  by  $\Delta a^s > 0$  and decrease  $a^c$  by  $\Delta a^c < 0$ ,  $V$  is unchanged and  $\underline{v}_f$  decreases by  $\Delta \underline{v}_f < 0$  as shown above. Hence if  $\pi^{*'}$  is enough small at  $a^s = 0$ , then the above incentive constraint is satisfied at  $\Delta a^s$  for  $(\Delta a^s, a^c + \Delta a^c, a^c + \Delta a^c, \dots)$ . This is a contradiction.

Suppose that  $\{\delta_2(t)\}_{t=1}^\infty$  is strictly more present biased relative to  $\{\delta_3(t)\}_{t=1}^\infty$  in the FOSD sense. If  $\underline{a}_f^\infty \neq (a^s, a^c, a^c, \dots)$ , then  $c_2(a^s, a^c, a^c, \dots) < c_2(\underline{a}_f^\infty)$ , a contradiction. Thus  $(a^s, a^c, a^c, \dots)$  is the unique worst forgiving punishment path in this case.  $\square$



## Proof of Proposition 7

*Proof.* By using a similar argument to the one in the proof of Proposition 6, we can show that there exists a worst equilibrium path that takes the form  $\underline{a}_{1,f}^\infty = (\widehat{a}^s, \widehat{a}^c, \widehat{a}^c, \dots)$  where  $\widehat{a}^s \leq \widehat{a}^c$ . Hence we can formulate the problems of finding the payoffs  $\underline{v}_f$ ,  $\underline{v}_{1,f}$ , and  $\bar{v}_{1,f}$  as follows:

$$\begin{aligned} \underline{v}_f &= \min_{a,a' \in A} \delta_2(1)\pi(a) + (1 - \delta_2(1))\pi(a') \\ \text{s.t. } & \frac{\delta(1)}{1 - \delta(1)}(\pi^*(a) - \pi(a)) \leq \delta_2(1)(\pi(a') - \pi(a)), \\ & \frac{\delta(1)}{1 - \delta(1)}(\pi^*(a') - \pi(a')) \leq \delta_2(1)(\pi(a') - \pi(a)), \\ & a \leq a', \end{aligned}$$

$$\begin{aligned} \underline{v}_{1,f} &= \min_{a,a' \in A} \delta(1)\pi(a) + (1 - \delta(1))\pi(a') \\ \text{s.t. } & \delta(1)(\pi^*(a) - \pi(a)) \leq (1 - \delta(1))(\pi(a') - \underline{v}_f), \\ & \delta(1)(\pi^*(a') - \pi(a')) \leq (1 - \delta(1))(\pi(a') - \underline{v}_f), \\ & a \leq a', \end{aligned}$$

and

$$\begin{aligned} \bar{v}_{1,f} &= \max_{a \in A} \pi(a) \\ \text{s.t. } & \delta(1)(\pi^*(a) - \pi(a)) \leq (1 - \delta(1))(\pi(a) - \underline{v}_f). \end{aligned}$$

The above problems depend only on  $\delta(1)$  and  $\delta_2(1)$  among  $\{\delta(t)\}_{t=1}^\infty$ , and thus the three payoffs are completely determined by  $\delta(1)$  and  $\delta_2(1)$ .

Choose any solution  $(a^s, a^c)$  to the problem of finding  $\underline{v}_f$  and any solution  $(\widehat{a}^s, \widehat{a}^c)$  to the problem of finding  $\underline{v}_{1,f}$ . A path in the form  $(a, a', a', \dots)$  is a forgiving equilibrium if and only if it is feasible for the problem of finding  $\underline{v}_{1,f}$ . Hence the constraint set of the problem is contained in the set  $A_c := \{(a, a') : \delta_2(1)\pi(a) + (1 - \delta_2(1))\pi(a') \geq \delta_2(1)\pi(a^s) + (1 - \delta_2(1))\pi(a^c)\}$ , while  $\delta(1)\pi(\widehat{a}^s) + (1 - \delta(1))\pi(\widehat{a}^c) \leq \delta(1)\pi(a^s) + (1 - \delta(1))\pi(a^c)$ . Since  $\delta(1) \geq \delta_2(1)$ , we have  $\delta(1)\pi(a) + (1 - \delta(1))\pi(a') > \delta(1)\pi(a^s) + (1 - \delta(1))\pi(a^c)$  for any  $(a, a') \in A_c$  such that  $a > a^s$ . Hence  $\widehat{a}^s \leq a^s$ . Then  $\widehat{a}^c \geq a^c$  in order to have  $(\widehat{a}^s, \widehat{a}^c) \in A_c$ .

For any  $a \in [\widehat{a}^s, \widehat{a}^c]$ , the path  $(a, \widehat{a}^c, \widehat{a}^c, \dots)$  is a forgiving equilibrium path, and thus we can achieve any payoff in  $[V(\underline{a}_{1,f}^\infty), \pi(\widehat{a}^c)]$  at equilibrium. Also, by Assumption 2,  $(a, a, \dots)$  is a forgiving equilibrium path for any  $a \in [\widehat{a}^c, a_f^o]$ , which shows that any payoff in

$[\pi(\hat{a}^c), \pi(a_f^o)]$  can be achieved at equilibrium. Hence the set of forgiving equilibrium payoff is  $[V(\underline{a}_{1,f}^\infty), \pi(a_f^o)]$ .

Consider the problem to obtain  $\underline{v}_f$ . Let  $(a^s, a^c)$  be a solution to the problem. Since  $a^s < a^c$  and at least one of the first two constraints should be binding at a solution, both sides of those constraints are positive. As we decrease  $\delta(1)$  while fixing  $\delta_2(1)$ , both constraints become slack at  $(a^s, a^c)$ , and thus we can decrease  $\underline{v}_f$  by reducing  $a^c$ . As we increase  $\delta_2(1)$  while fixing  $\delta(1)$ ,  $(a^s, a^c)$  is still feasible while the objective value becomes smaller at  $(a^s, a^c)$ . Thus the change decreases  $\underline{v}_f$ . Next consider the problem to obtain  $\underline{v}_{1,f}$ . Let  $(\hat{a}^s, \hat{a}^c)$  be a solution to the problem. As we increase  $\delta_2(1)$ ,  $\underline{v}_f$  decreases and thus the first two constraints become slack at  $(\hat{a}^s, \hat{a}^c)$ . Then we can decrease  $\underline{v}_{1,f}$  by reducing  $\hat{a}^s$  or  $\hat{a}^c$ . Lastly consider the problem to obtain  $\bar{v}_{1,f}$ . Let  $a_f^o$  be a solution to the problem. As we decrease  $\delta(1)$  or increase  $\delta_2(1)$ ,  $\underline{v}_f$  decreases and thus the incentive constraint becomes slack at  $a_f^o$ . Hence we can increase  $\bar{v}_{1,f}$  until it reaches  $\bar{\pi}$ .  $\square$

### Proof of Proposition 8

*Proof.* Suppose that the  $\{\beta_t\}_t$ -weighted discounting is present biased in the MLRP sense. Then  $\frac{\beta_t}{\beta_{t-1}} \leq \frac{\beta_{t+1}}{\beta_t}$  for all  $t = 1, 2, \dots$ . The discounting is a geometric discounting (with discount factor  $\beta\delta$ ) if and only if  $\frac{\beta_t}{\beta_{t-1}}$  is some constant  $\beta > 0$  for all  $t = 1, 2, \dots$ . Since the discounting is not geometric, we have  $\frac{\beta_t}{\beta_{t-1}} < \frac{\beta_{t+1}}{\beta_t}$  for some  $t$ . Suppose that  $\beta_1 \geq 1$ . Then  $1 \leq \frac{\beta_1}{\beta_0} \leq \frac{\beta_2}{\beta_1} \leq \frac{\beta_3}{\beta_2} \leq \dots$  with at least one inequality strict. Then  $\{\beta_t\}$  is unbounded, which is a contradiction. Thus  $\beta_1 < 1$ . If  $\sum_{t=1}^\infty \beta_t = \infty$ , then  $C(\{\beta_t\}_t) = \frac{1}{\beta_1} > 1$ . Suppose that  $\sum_{t=1}^\infty \beta_t < \infty$ . By the MLRP assumption, we have  $\frac{\beta_{1+k}}{\beta_1} \geq \frac{\beta_k}{\beta_0}$  for all  $k = 1, 2, \dots$ , and since the discounting is not geometric, we have strict inequality for some  $k$ . This implies  $\frac{\beta_2}{\beta_1} + \frac{\beta_3}{\beta_1} + \frac{\beta_4}{\beta_1} + \dots > \frac{\beta_1}{\beta_0} + \frac{\beta_2}{\beta_0} + \frac{\beta_3}{\beta_0} + \dots$ , or  $\sum_{t=2}^\infty \frac{\beta_t}{\beta_1} > \beta_1$ . Hence we have  $C(\{\beta_t\}_t) > 1$ .

Since  $C(\{\beta_t\}_t) > 1$ , the objective function of the problem to define  $a^{s*}$  is strictly convex by Assumption 2. Since the objective function is continuous and  $A$  is compact, there exists a unique solution  $a^{s*}$ , which implies the uniqueness of  $a^{c*}$ . Note that, for any constant  $k > 1$ ,  $\pi(a) + k(\pi^*(a) - \pi(a)) > 0$  for all  $a \in A$ . Hence  $\pi(a^{s*}) + C(\{\beta_t\}_t)(\pi^*(a^{s*}) - \pi(a^{s*})) > 0$  and  $\pi(a^{c*}) > 0$ .

Consider a stick-and-carrot path  $\underline{a}_f^\infty = (a^s, a^c, a^c, \dots)$  that achieves  $\underline{v}_f$ . We know that  $a^s < a^c$ . We also know that  $a^c < a_f^o = a^*$  when  $\delta$  is large. Thus the incentive constraint is binding at  $a^s$  only. Then we can obtain  $\underline{v}_f$  just by minimizing  $c_2(\underline{a}_f^\infty) = \delta_2(1)\pi(a^s) + (1 - \delta_2(1))\pi(a^c)$  subject to the incentive constraint in the first period:

$$\delta(1)(\pi^*(a^s) - \pi(a^s)) = \delta(2)(\pi(a^c) - \pi(a^s)).$$

For  $\{\beta_t\}_t$ -weighted discounting, this problem can be written as

$$\begin{aligned} \underline{v}_f &= \min_{a^c, a^s \in A} \frac{1}{\sum_{t=1}^{\infty} \beta_t \delta^t} \left( \beta_1 \delta \pi(a^s) + \sum_{t=2}^{\infty} \beta_t \delta^t \pi(a^c) \right) \\ \text{s.t. } \pi(a^c) &= \pi(a^s) + \frac{1}{\beta_1 \delta} (\pi^*(a^s) - \pi(a^s)). \end{aligned}$$

Substituting the constraint into the objective function, this problem can be rewritten as

$$\underline{v}_f = \min_{a^s \in [0, a^*]} \pi(a^s) + \frac{(1-\delta) \sum_{t=2}^{\infty} \beta_t \delta^t}{(1-\delta) \sum_{t=1}^{\infty} \beta_t \delta^t} \frac{1}{\beta_1 \delta} (\pi^*(a^s) - \pi(a^s)).$$

As  $\delta \rightarrow 1$ , the coefficient of  $(\pi^*(a^s) - \pi(a^s))$  in the objective function converges to  $C(\{\beta_t\}_t)$ . Since the solution of the above problem in the limit as  $\delta \rightarrow 1$  is given by the solution of the limit of the problem, the stick and carrot path must converge to  $(a^{s*}, a^{c*}, a^{c*}, \dots)$ .

Clearly  $\underline{v}_f$  converges to  $\pi(a^{s*}) + C(\{\beta_t\}_t) (\pi^*(a^{s*}) - \pi(a^{s*}))$  as  $\delta \rightarrow 1$ , which coincides with  $\pi(a^{c*})$  if  $\sum_{t=1}^{\infty} \beta_t = \infty$ .  $\square$

### Proof of Proposition 9

*Proof.* Suppose that  $\underline{a}^{t, \infty}$  is a nonconstant forgiving path for some  $t$ . Then we can find  $a^c \in (\underline{a}^t, a^o]$  such that  $V(\underline{a}^{1,t}, a^c, a^c, \dots) = V(\underline{a}^{\infty})$ . This is equivalent to  $c_{t+1}(a^c, a^c, \dots) = c_{t+1}(\underline{a}^{t+1, \infty})$ , which implies that  $c_{\tau}(a^c, a^c, \dots) \geq c_{\tau}(\underline{a}^{t+1, \infty})$  for every  $\tau = 1, 2, \dots, t$ . Thus the incentive constraints up to period  $t$  are still satisfied for the new path  $(\underline{a}^{1,t}, a^c, a^c, \dots)$ . We can show that the incentive to play  $a^c$  is satisfied as in the proof of Proposition 6. So  $(\underline{a}^{1,t}, a^c, a^c, \dots)$  is an ASPE path. Since  $c_{t+2}(a^c, a^c, \dots) \leq c_{t+2}(\underline{a}^{t+1, \infty})$ , we have  $c_2(\underline{a}^{1,t}, a^c, a^c, \dots) \leq c_2(\underline{a}^{\infty}) = \underline{v}$ , and thus  $(\underline{a}^{1,t}, a^c, a^c, \dots)$  is a worst punishment path as well.

We consider two cases. First suppose that  $a^c$  is not the best equilibrium action  $a^o$ . Since  $a^c > \underline{a}^t$ , the incentive constraint to play  $a^c$  is not binding. Hence we can find  $a'$  and  $a''$  in a neighborhood of  $a^c$  such that  $a' > a^c > a''$  and  $c_{t+1}(a^c, a^c, \dots) = c_{t+1}(a', a'', a'', \dots)$ . Since  $a'$  and  $a''$  are close to  $a^c$ , the incentive constraints in periods  $t+1$  and  $t+2$  are not violated. The incentive constraints in all the periods up to period  $t$  are satisfied as before. Hence the path  $(\underline{a}^{1,t}, a', a'', a'', \dots)$  is an ASPE path. Since  $\delta_{\tau}(1)$  is decreasing in  $\tau$  (in particular,  $\delta_{t+1}(1) > \delta_{t+2}(1)$ ), we have  $c_2(\underline{a}^{1,t}, a^c, a^c, \dots) > c_2(\underline{a}^{1,t}, a', a'', a'', \dots)$ , which is a contradiction.

Next suppose that  $a^c = a^o$ . Then we can increase  $\underline{a}^t$  into  $a'$  and decrease  $a^c$  into  $a''$  slightly so that  $V(\underline{a}^{1,t-1}, a', a'', a'', \dots) = V(\underline{a}^{1,t}, a^c, a^c, \dots)$ . We show that this new path  $(\underline{a}^{1,t-1}, a', a'', a'', \dots)$  is an ASPE path. The incentive constraints up to period  $t-1$  are satisfied as before.  $(a'', a'', \dots)$  is clearly an ASPE path. If  $\delta(1)$  is small enough, then we can make the incentive constraint slack at  $a$  for a path  $(a, a^o, a^o, \dots)$  for any

$a \in [0, a^*]$  (in which case  $a^o = a^*$  must hold). Hence the incentive constraint at period  $t$  holds after this perturbation as long as  $\delta(1)$  is small enough. However,  $c_2(\underline{a}^{1,t-1}, a', a'', a'', \dots) < c_2(\underline{a}^{1,t}, a^c, a^c, \dots)$  holds because  $\delta_\tau(1)$  is decreasing in  $\tau$  (in particular,  $\delta_t(1) > \delta_{t+1}(1)$ ). This is a contradiction. Hence if  $\underline{a}^{t,\infty}$  is a forgiving path for some  $t$ , then it must be a constant path. Furthermore, if this constant path is not  $(\widehat{a}, \widehat{a}, \dots)$ , then either it is  $(a^o, a^o, \dots)$ , which is obviously not the worst punishment path, or it is a constant path with an unbinding incentive constraint by Assumption 2, in which case we can do a similar perturbation to create a worse punishment path. Hence this path must be  $(\widehat{a}, \widehat{a}, \dots)$ .

We already know that no constant path can be a worst punishment path, so  $t$  cannot be 1. If  $\underline{a}^{2,\infty}$  is a constant path, then  $\underline{a}^1 > \underline{a}^2$  must be the case as  $\underline{a}^\infty$  should not be a forgiving punishment path. But then  $\underline{a}^{2,\infty}$  would generate a worse punishment payoff than  $\underline{a}^\infty$ , which is a contradiction. Hence  $t$  cannot be 2 either.  $\square$

### Proof of Proposition 10

*Proof.* Suppose that the graph of  $W$  is nonempty and convex. For any  $\pi^{1,T-1}$  such that  $W(\pi^{1,T-1})$  is nonempty,  $W'(\pi^{NE}, \pi^{1,T-2})$  is nonempty, and thus the graph of  $W'$  is nonempty. Choose any  $(\pi_1^{1,T-1}, v_1)$  and  $(\pi_2^{1,T-1}, v_2)$  in the graph of  $W'$  and any  $\alpha \in [0, 1]$ . Let  $\pi_\alpha^{1,T-1} = \alpha\pi_1^{1,T-1} + (1-\alpha)\pi_2^{1,T-1}$  and  $v_\alpha = \alpha v_1 + (1-\alpha)v_2$ . Then it suffices to show that  $v_\alpha \in W'(\pi_\alpha^{1,T-1})$ . Let  $(\pi'_i, v'_i)$  be a pair such that  $(1-\delta)\pi'_i + \delta v'_i = v_i$  while satisfying all the constraints in (3) to define  $W'(\pi_i^{1,T-1})$ , for  $i = 1, 2$ . Let  $\pi'_\alpha = \alpha\pi'_1 + (1-\alpha)\pi'_2$  and  $v'_\alpha = \alpha v'_1 + (1-\alpha)v'_2$ . Since  $\pi^d(\pi)$  is convex and the graph of  $W$  is convex,  $(\pi'_\alpha, v'_\alpha)$  satisfies all the constraints to define  $W'(\pi_\alpha^{1,T-1})$  while  $(1-\delta)\pi'_\alpha + \delta v'_\alpha = v_\alpha$ . Hence  $v_\alpha \in W'(\pi_\alpha^{1,T-1})$ .  $\square$

### Proof of Proposition 11

*Proof.* Note that  $(\pi', v')$  is a solution to

$$\begin{aligned} & \min_{\pi, v} (1-\delta)\pi + \delta v \\ & \text{s.t. } \delta_2(1)\pi^2 + \dots + \delta_2(T-2)\pi^{T-1} \\ & \quad + [1 - \delta_2(1) - \dots - \delta_2(T-2)][\delta_T(1)\pi + (1-\delta_T(1))v] \\ & \geq \underline{v} + \frac{\delta(1)}{1-\delta(1)}[\pi^d(\pi^1) - \pi^1], \\ & \quad \underline{\pi} \leq \pi \leq \bar{\pi}, v \in W^*(\pi^{2,T-1}, \pi). \end{aligned}$$

Suppose that  $v' > \underline{W}^*(\pi^{2,T-1}, \pi')$  and  $\pi' \neq \bar{\pi}$ . If  $\pi' = \bar{v}$ , then  $W^*(\pi^{2,T-1}, \pi')$  must be a singleton  $\{\bar{v}\}$  because  $\bar{v}$  can be supported only by a constant sequence  $(\bar{v}, \bar{v}, \dots)$ . This contradicts  $v' > \underline{W}^*(\pi^{2,T-1}, \pi')$ , hence  $\pi' < \bar{v}$ . Then, since  $\underline{W}^*(\pi^{2,T-1}, \pi)$  is continuous in the last argument, we can increase  $\pi'$  and decrease  $v$  so that  $\delta_T(1)\pi + (1-\delta_T(1))v$  remains

the same and all the constraints are still satisfied. Since  $\frac{\delta_T(1)}{1-\delta_T(1)} > \frac{1-\delta}{\delta}$ , this change makes the objective value smaller, a contradiction. This proves the first result.

Take any worst punishment path  $\underline{\pi}^\infty$  such that  $a^t < a^*$  for every  $t$ . Let  $(\underline{\pi}^{1,T-1}, v^T)$  be the associated payoffs and the continuation score. Note that  $(\underline{\pi}^{1,T-1}, v^T)$  is a solution to

$$\begin{aligned} \underline{v} &= \min_{\pi^{1,T-1}, v} \delta_2(1)\pi^1 + \dots + \delta_2(T-1)\pi^{T-1} + [1 - \delta_2(1) - \dots - \delta_2(T-1)]v \\ \text{s.t. } &\pi^{1,T-1} \in [\underline{\pi}, \bar{\pi}]^{T-1}, v \in W^*(\pi^{1,T-1}). \end{aligned}$$

If  $v^T > \underline{W}^*(\underline{\pi}^{1,T-1})$ , we can reduce the objective value by reducing  $v^T$ . Hence,  $v^T = \underline{W}^*(\underline{\pi}^{1,T-1})$ . Then we can apply the previous result recursively to obtain  $v^{t+T-1} = \underline{W}^*(\underline{\pi}^{t+1, T+t-2}, \pi^{t+T-1})$  for  $t = 1, 2, \dots$ , since  $\pi^{t+T-1} < \bar{\pi}$  for every  $t$  by assumption.

Finally, take any  $(\underline{\pi}^{1,T-1}, v^T)$  that solves (2) with respect to  $W^*$ . Consider the following set

$$\{(\pi, v) | (1-\delta)\pi + \delta v = v^T \ \& \ v \in W^*(\underline{\pi}^{2,T-1}, \pi)\}.$$

Not all  $(\pi, v)$  in this set may satisfy the one-shot deviation constraint for the first period. But, there exist some  $(\pi, v)$  that does, because  $v^T \in W^*(\underline{\pi}^{1,T-1})$ . Since  $\delta_T(1) > 1 - \delta$  (by  $\beta_{T-1} > \beta_T$ ), if the incentive constraint is satisfied for  $(\pi, v)$ , then it must be satisfied for any  $(\pi', v')$  such that  $\pi' > \pi$  and  $(1-\delta)\pi' + \delta v' = v^T$ . Hence if we take  $(\pi', v')$  in the above set where  $\pi'$  is largest, then the one shot deviation constraint for the first period must be satisfied for such  $(\pi', v')$ . Define  $(\underline{\pi}^T, v^{T+1})$  as such a point in the above set. We can repeat this to construct an entire sequence of payoffs  $\underline{\pi}^\infty$  that generate  $\underline{v}$  as the continuation payoff. Since all the one shot deviation constraints are satisfied by construction, the associated action path  $a^\infty$  is an ASPE path, hence a worst punishment path.  $\square$

### Proof of Proposition 12

*Proof.* Let  $\underline{a}^\infty$  be a worst punishment path. With  $\beta$ - $\delta$  discounting, we have  $c_2 = c_3 = \dots$ , and let us denote them by  $c$ . Since  $\underline{a}^{2,\infty}$  is an equilibrium path, we have  $c(\underline{a}^{2,\infty}) \leq \pi(a^o)$ . Thus, there exists unique  $a^c \in [0, a^o]$  such that  $\pi(a^c) = c(\underline{a}^{2,\infty})$ . Let  $a^s = \underline{a}^1$ . By construction, we have  $c(a^s, a^c, a^c, \dots) = c(\underline{a}^\infty) = \underline{v}$ . Note that  $a^s \leq a^c$  because if  $a^s > a^c$  then  $c(\underline{a}^{2,\infty}) < c(a^s, a^c, a^c, \dots) = \underline{v}$ , a contradiction. We can show that  $(a^s, a^c, a^c, \dots)$  is an equilibrium path and  $a^s > a^c$ , as in the proof of Proposition 6-1.  $\square$

### Proof of Proposition 13

*Proof.* Since the preference exhibits future bias in the MLRP sense, by Proposition 3 we obtain the worst punishment (payoff) path  $(\underline{\pi}, \dots, \underline{\pi}, \pi, \bar{\pi}, \bar{\pi}, \dots)$  in stick-and-carrot form

where  $\underline{\pi}$  is played  $K$  times.  $K$  increases to  $\infty$  as  $\Delta$  gets smaller, and thus we have  $K \geq 1$  when  $\Delta$  is small enough. Since the incentive constraint for the first period must be binding, it is given by

$$\delta(1)(0 - \underline{\pi}) = \delta(K + 1)(\pi - \underline{\pi}) + \delta(K + 2)(\bar{\pi} - \pi).$$

The limit of the incentive constraint as  $\Delta$  goes to zero can be written as

$$(1 - \eta(0))(-\underline{\pi}) = (1 - \eta(\hat{T}))e^{-r(\hat{T} - \int_0^{\hat{T}} \eta(s)ds)}(\bar{\pi} - \underline{\pi}), \quad (6)$$

where  $\hat{T}$  denotes the limit of  $K\Delta$ .  $(1 - \eta(\hat{T}))e^{-r(\hat{T} - \int_0^{\hat{T}} \eta(s)ds)}$  is equal to  $1 - \eta(0)$  at  $\hat{T} = 0$ , and it may increase for small  $\hat{T}$  but it decreases eventually. Thus,  $\hat{T}$  is determined uniquely by the above incentive constraint, and it represents the time around which there is a switch from  $\underline{\pi}$  to  $\bar{\pi}$ .

The worst punishment payoff in the limit as  $\Delta$  goes to zero is the integral of  $\underline{\pi}$  from 0 to  $\hat{T}$  plus that of  $\bar{\pi}$  from  $\hat{T}$  to  $\infty$  with respect to the discounting function  $\rho(\tau)$ . That is,

$$\lim_{\Delta \rightarrow 0} \underline{v} = \left( \int_0^{\hat{T}} \rho(\tau) d\tau \right) \underline{\pi} + \left( \int_{\hat{T}}^{\infty} \rho(\tau) d\tau \right) \bar{\pi}.$$

Using the relationship (6), we obtain

$$\begin{aligned} \lim_{\Delta \rightarrow 0} \underline{v} &= \left[ -e^{-r(\tau - \int_0^{\tau} \eta(s)ds)} \right]_0^{\hat{T}} \underline{\pi} + \left[ -e^{-r(\tau - \int_0^{\tau} \eta(s)ds)} \right]_{\hat{T}}^{\infty} \bar{\pi} \\ &= (1 - e^{-r(\hat{T} - \int_0^{\hat{T}} \eta(s)ds)}) \underline{\pi} + e^{-r(\hat{T} - \int_0^{\hat{T}} \eta(s)ds)} \bar{\pi} \\ &= \underline{\pi} + e^{-r(\hat{T} - \int_0^{\hat{T}} \eta(s)ds)} (\bar{\pi} - \underline{\pi}) \\ &= \underline{\pi} - \frac{1 - \eta(0)}{1 - \eta(\hat{T})} \underline{\pi} = \frac{\eta(0) - \eta(\hat{T})}{1 - \eta(\hat{T})} \underline{\pi}. \end{aligned}$$

□

## Proof of Corollary 2

*Proof.* Choose any  $v \in (\frac{\eta(0) - \eta(\hat{T})}{1 - \eta(\hat{T})} \underline{\pi}, \bar{\pi}]$ . Since  $\frac{\eta(0) - \eta(\hat{T})}{1 - \eta(\hat{T})} \underline{\pi} > \underline{\pi}$ , there exists  $a \in A$  such that  $\pi(a) = v$ . Consider the constant path  $a^\infty = (a, a, \dots)$ . It is an equilibrium path if the following incentive constraint is satisfied:

$$\pi^*(a) - \pi(a) \leq \frac{1 - \delta(1)}{\delta(1)} (\pi(a) - \underline{v}).$$

As  $\Delta \rightarrow 0$ ,  $\delta(1) \rightarrow 0$  and  $\underline{v} \rightarrow \frac{\eta(0) - \eta(\hat{T})}{1 - \eta(\hat{T})} \underline{\pi}$ . Thus the incentive constraint must be satisfied

in the limit. □

### Proof of Proposition 14

*Proof.* Let  $\underline{a}^\infty$  be the worst punishment path that yields the payoff stream  $(\underline{\pi}, \dots, \underline{\pi}, \pi, \bar{\pi}, \bar{\pi}, \dots)$ , where  $\underline{a}$  is played  $K$  times. When  $r$  is sufficiently small, we have  $K \geq T$ . Following the proof of Proposition 4, we can obtain

$$\underline{v} = \frac{e^{-r(\Delta - \int_0^\Delta \eta(s) ds)} - e^{-r\Delta}}{e^{-r(\Delta - \int_0^\Delta \eta(s) ds)} - e^{-r(2\Delta - \int_0^\Delta \eta(s) ds)}} \underline{\pi}.$$

Hence,

$$\lim_{r \rightarrow 0} \underline{v} = \frac{1}{\Delta} \left( \int_0^\Delta \eta(s) ds \right) \underline{\pi}.$$

□

## References

- [1] D. Abreu, “Extremal Equilibria of Oligopolistic Supergames,” *Journal of Economic Theory* 39 (1986) 191–225.
- [2] D. Abreu, D. Pearce, and E. Stacchetti, “Toward a Theory of Discounted Repeated Games with Imperfect Monitoring,” *Econometrica* 58 (1990) 1041–1063.
- [3] H. Chade, P. Prokopovych, and L. Smith, “Repeated Games with Present-Biased Preferences,” *Journal of Economic Theory* 139 (2008) 157–175.
- [4] S. Frederick, G. Loewenstein, and T. O’Donoghue, “Time Discounting and Time Preference: A Critical Review,” *Journal of Economic Literature* XL (2002) 351–401.
- [5] A. Kochov and Y. Song, “Repeated Games with Endogenous Discounting,” working paper (2015).
- [6] D. Laibson, “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics* 112 (1997) 443–477.
- [7] D. Laibson, “Life Cycle Consumption and Hyperbolic Discount Functions,” *European Economic Review* 42 (1998) 861–871.
- [8] I. Obara and J. Park, “Repeated Games with General Time Preference,” working paper (2014).

- [9] T. O'Donoghue and M. Rabin, "Doing it Now or Later," *American Economic Review* 89 (1999) 103–124.
- [10] B. Peleg and M. Yaari, "The Existence of a Consistent Course of Action when Tastes are Changing," *Review of Economic Studies* 40 (1973) 391–401.
- [11] E. Phelps and R. Pollak, "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies* 35 (1968) 185–199.