

Equilibrium Restoration in a Class of Tolerant Strategies

by Romeo Balanquit
Centre for Economic Studies and Planning
Jawaharlal Nehru University
New Delhi, India

Abstract

This study shows that in a two-player infinitely repeated game where one is impatient, Pareto-superior subgame perfect equilibria can still be achieved. An impatient player in this paper is depicted as someone who can truly destroy the possibility of attaining any feasible and individually rational outcome that is supported in equilibrium in repeated games, as asserted by the Folk Theorem. In this scenario, the main ingredient for the restoration of equilibrium is to introduce the notion of tolerant trigger strategy. Consequently, the use of the typical trigger strategy is abandoned since it ceases to be efficient as it only brings automatically the game to its punishment path, therefore eliminating the possibility of extracting other feasible equilibria. I provide a simple characterization of perfect equilibrium payoffs under this scenario and show that cooperative outcome can be approximated.

An eye for an eye makes the whole world blind.
- Mahatma Gandhi

1. Introduction

A central requirement in obtaining a cooperative outcome in repeated games, apart from the inherent need of a credible threat of punishment, is that players must be sufficiently patient. Future payoffs need to be valued highly so as not to induce anyone to deviate from any long-term contract. Otherwise, even those who supposed to reprove a deviant may also find it less attractive to impose punishment if the payoff for doing so decreases rapidly through time. Thus, it reaffirms why

only minimal amount of discounting is permitted so that the Folk Theorem can still be maintained¹.

When a certain player in a game is impatient, it is easy to see that any co-operative effort is hardly sustainable through time for that player wishes only to extract the highest gain the soonest possible. Such player does not even have to worry about future punishments since the future is less meaningful to him. So despite the good intentions others may have in leading the game to better results, their knowledge of the presence of the impatient player compels them not to attempt for any risky cooperative action from the very start. Thus, the game simply reverts to the unwanted equilibrium of the original single-stage game. This is perhaps one reason why in the literature of repeated games, studying a game involving an impatient player does not get so much attention. In this paper, I address the prospect of restoring some Pareto-superior equilibria by adopting a different strategy in such scenario where an impatient player is involved.

Imagine a two-player infinitely-repeated game where players have different discount factors or, say simply, different temperaments: one is patient and the other is impatient. Suppose further that the impatient one is so impatient that even the harshest punishment of penalizing him forever, in case he deviates from the typical trigger strategy², would not matter to him since he only cares for the current period. He therefore cannot be trusted to cooperate since cheating in the first period is always more rewarding to him. On the other end, as the patient player is aware of this, she may simply apply at the outset a strategy that will minimize her opponent's maximum payoff (i.e. minimax strategy) thus, eliminating any possibility of achieving a cooperative outcome.

One can argue, however, that the end of this game depends so much on how the

¹The Folk Theorem asserts that all feasible and individually rational outcomes (i.e. outcomes that are Pareto-superior to the minimax payoff of the initial stage game) can be supported in equilibrium in an infinitely-repeated game. Aumann & Shapley (1976) and Rubinstein (1979) showed that this set of equilibria is in fact subgame perfect. Abreu (1988) and Fudenberg & Maskin (1986) later showed that this is also true when minimal discounting on future payoffs is applied.

²The typical trigger strategy referred here is when both players continue to play the cooperative outcome for as long as no one has deviated in the past. In case either player deviates, both respond by defecting forever.

patient player will play the game. Being a patient person, she has the capacity to tolerate the other player initially, even at the expense of getting a very low payoff, provided that this gesture makes the other player cooperate in the succeeding stages. In this paper, I show that this set-up is subgame-perfect and that both players are made better-off than in a situation when no such tolerance is initiated.

Lehrer and Pauzner (1999) examined this case of unequal discounting between two players, although they maintained that these players remain *very patient on the absolute scale*. The expanded frontiers of the feasible set and the shape of the set of perfect equilibria that they have pinned down are therefore not the ones that will be obtained in the presence of an impatient player. The cooperative outcome may no longer be attainable and in some cases of very low discount factor, the set of feasible and (strictly) individually rational payoffs within the typical convex hull can be totally annihilated³. Certainly, Folk Theorem can no longer be aspired, nonetheless we show that certain class of tolerant strategies which uses the disparity in the time preferences of the two players can restore some perfect equilibria that are Pareto-superior to the stage-game Nash.

The notion of heterogenous discounting has clearly given greater possibilities for generating perfect equilibrium outcomes as shown also in a related two-player model of Salonen and Vartiainen (2008) and in the n -player setup proposed by Chen (2007). A more general result by Gueron *et. al.* (2010) even shows that any individually rational payoff that is below and thus nullified by the effective minimax value (a concept introduced by Wen (1994) for n -player games) can be restored in equilibrium. While these studies make use of unequal discounting, all of them maintain that the players' discount factors are sufficiently close to one. By the presence of an impatient player, this paper is therefore distinguished from these studies although it is not extended to n -player games as will be discussed in the final section.

Generally, the structure of a tolerant strategy along its (initial) contract path is a deterministic sequence of pure-strategy actions. In particular, we study those

³In Theorem 1, this very low discount factor is given a lower bound.

types that exhibit periodic structure over time under a perfect monitoring environment. For example, a patient player may agree to tolerate the other for two stages provided that a cooperative play is performed in the next three stages, and then tolerate again for the next two stages, and so on. This cyclical set-up works continuously *ad infinitum* for as long as no deviation has occurred in the past. A deviation at any time from either player leads the game to its punishment phase that imposes minimax strategies. We assume in this paper that such strategies are observable if these can only be implemented through mixed-strategy actions.

When a cooperative outcome is not attained, it is true that some correlated strategies between the two players could still approximate it despite having a reduced set of equilibrium payoffs caused by the impatient player. However, employing tolerant strategies do no less. It can further be shown that even in an extreme case of “impatience” (see Section 6), when public randomization can no longer generate individually rational equilibrium payoffs under the normal trigger strategy, these tolerant trigger strategies can still continue to generate some of these equilibria.

The next section illustrates the main idea of this paper through a concrete example. Section 3 establishes the environment governing around the problem while sections 4-7 provide a formal analysis. Section 8 concludes by discussing some difficulties in generalizing some results.

2. Example

Consider a Prisoner’s Dilemma game with the following payoffs:

	C	D
C	3,3	0,4
D	4,0	1,1

The minimax point of this game is (1,1) and for $\delta < \frac{1}{3}$, an infinitely repeated game cannot obtain any equilibrium other than the players’ minimax point, which in this case is also a Nash equilibrium. Hence, each player will only settle to receive an average payoff of 1 in the repeated game.

Assume now that the two players have different valuation of time: $\delta_1 < \frac{1}{3}$ and $\delta_2 \geq \frac{1}{3}$. Then, suppose that Player 2 offers a strategy wherein she will always play C provided that Player 1 alternates his actions between D and C, starting with D. Any deviation from this strategy from either player prompts both of them to play D forever after. In other words, Player 2 tolerates Player 1 in stage one (and in all succeeding odd-number stages) and endures receiving 0, which is even lower than her minimax payoff.

The rationale behind Player 2's offer is that if this strategy succeeds, she will receive an average income⁴ of $\frac{3\delta_2}{1+\delta_2}$ (*i.e.* $0 + 3\delta_2 + 0 + 3\delta_2^3 + \dots$) which is greater than her average income when no such offer is made, provided that $\delta_2 > \frac{1}{2}$. On the part of Player 1, he will accept the offer since this strategy promises him an average payoff of $\frac{4+3\delta_1}{1+\delta_1}$ (*i.e.* $4 + 3\delta_1 + 4\delta_1^2 + 3\delta_1^3 + \dots$), which is always a lot more than what he will get when he is not tolerated.

This strategy is a subgame perfect equilibrium and is shown in the following manner. Observe that Player 1 will not think of deviating from playing D in the 1st stage knowing that he will be tolerated by Player 2. If he were to think of deviating, it must be in the 2nd stage where he is bound to get a lower payoff by reciprocating Player 2's goodwill. Deviating in the 2nd stage therefore becomes irresistible when his average income from the path $(4, 4\delta_1, \delta_1^2, \delta_1^3, \dots)$ exceeds that of simply sticking to the strategy, *i.e.* $(4, 3\delta_1, 4\delta_1^2, 3\delta_1^3, \dots)$. This condition is presented as:

$$(1 - \delta_1) \left(4 + 4\delta_1 + \frac{\delta_1^2}{1 - \delta_1} \right) > (1 - \delta_1) \left(\frac{4 + 3\delta_1}{1 - \delta_1^2} \right) \Rightarrow 3\delta_1^2 + 3\delta_1 - 1 < 0$$

Solving for δ_1 , Player 1 will deviate when $\delta_1 < \frac{\sqrt{21}-3}{6} \approx 0.26$.

For Player 2, deviating in the 1st stage, *i.e.* playing D, will only bring back the game to its minimax point which means that both players ended up playing (D,D) in every stage thereafter. Besides, she would not opt to deviate at this stage

⁴Average income is computed in its discounted form over infinite stages as $(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} P^t$, where P^t is the payoff at stage t . Note also that the formula $1 + \delta + \delta^2 + \dots + \delta^{n-1} = \frac{1-\delta^n}{1-\delta}$ will be extensively used in this paper.

knowing that her offer will be rewarding in the long run, for as long as $\delta_2 > \frac{1}{2}$. The case is different in the 2nd stage where there arises also a temptation for her to deviate. This possible deviation is realized when the path $(0, 4\delta_2, \delta_2^2, \delta_2^3, \dots)$ becomes more profitable than $(0, 3\delta_2, 0, 3\delta_2^3, \dots)$. That is:

$$(1 - \delta_2) \left(4\delta_2 + \frac{\delta_2^2}{1 - \delta_2} \right) > (1 - \delta_2) \left(\frac{3\delta_2}{1 - \delta_2^2} \right) \Rightarrow 3\delta_2^2 - \delta_2 - 1 < 0$$

Solving for δ_2 , Player 2 will deviate when $\delta_2 < \frac{\sqrt{13}+1}{6} \approx 0.77$.

One can check that the condition for deviating in all subsequent odd-number stages of the repeated game is similar to the respective condition each player face during the 1st stage. Similarly, all succeeding even-number stages establish the same condition as in the 2nd stage, respective to each player (see Lemma 1). Thus, the Nash equilibrium of this infinitely repeated game exists for:

$$\delta_1 \in \left[\frac{\sqrt{21}-3}{6}, \frac{1}{3} \right) \quad \text{and} \quad \delta_2 \in \left[\frac{\sqrt{13}+1}{6}, 1 \right).$$

When either player deviates at any stage of the game, the strategy calls for each player a punishment path of minimaxing each other thereafter *i.e.* playing (D,D) from then on. And since (D,D) is a Nash equilibrium of the prisoner's dilemma game, neither player can gain by deviating from this punishment path which establishes credibility in rendering punishment⁵. Therefore, the set of Nash equilibrium points of this strategy also satisfies subgame perfection.

By inputting all the equilibrium-generating values of δ_1 and δ_2 into the average income of each player, the set of all possible perfect equilibrium payoffs is illustrated in Figure 1 as a rectangular block. Notice that it is outside the typical feasible and individually rational set of payoffs generated in a repeated game with very patient players.

⁵In the example of prisoner's dilemma, the minimax punishment inherently coincides with the Cournot-Nash reversion, extensively used by Friedman (1971). A more general minimaxing punishment scheme is presented in Section 7 for any two-person game.

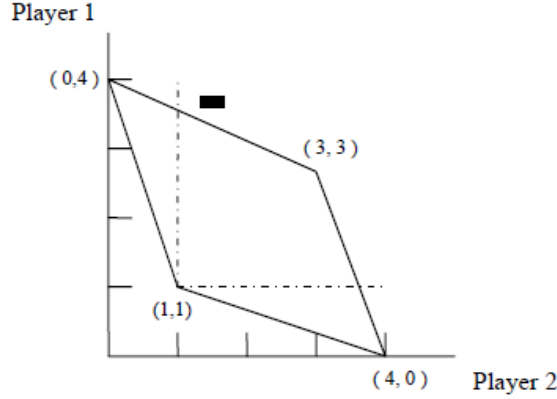


Figure 1: In the above example where cooperative outcome is no longer attainable, the set of equilibrium payoffs is depicted by the rectangular block that is outside the typical convex hull.

3. Framework

Consider an ∞ -fold repeated game $\Gamma^\infty(\delta_M, \delta_P)$ with two players, M (impatient) and P (patient), and their respective discount factors δ_M and δ_P , where $0 < \delta_M < \delta_P < 1$. Write $(a_M, a_P) \in S_M \times S_P$ as a vector of outcomes/actions within the *pure* strategy space and $\mu_i(a_M, a_P) : S_M \times S_P \rightarrow \mathbb{R}$ as the continuous payoff function of i , where $i = M, P$. For convenience, we denote an unsubscripted bold symbol as a vector of two players (*e.g.* $\mathbf{a} = (a_M, a_P)$) and denote $-i$ to refer to the other player. Fix the minimax payoff for each player as $\hat{V}_i = \min_{a_{-i} \in S_{-i}} \max_{a_i \in S_i} \mu_i(a_i, a_{-i})$. For a set of feasible payoffs F , which is also defined as the convex hull of the set $\{(V_M, V_P) \mid \boldsymbol{\mu}(\mathbf{a}) = (V_M, V_P), \text{ for some } \mathbf{a} \in S_M \times S_P\}$, denote a subset R to be the set of individually rational outcomes *i.e.* $R = \{(V_M, V_P) \in F \mid V_i > \hat{V}_i, \text{ for both } i\}$. Let (\bar{V}, \underline{V}) be some feasible vector payoff not in R since $\underline{V} < \hat{V}_i < V_i < \bar{V}$ for both i , where \bar{V} being the highest possible payoff to i and \underline{V} being the lowest. In this symmetric two-person game, we set $\hat{V}_i = \hat{V}$ and the cooperative outcome as $(V_c, V_c) \in R$. Finally, we assume that $\Gamma^\infty(\delta_M, \delta_P)$ is played under a complete information environment.

The game $\Gamma^\infty(\delta_M, \delta_P)$ is played throughout a discrete time denoted by $t \in \{1, 2, 3, \dots\}$. Let its outcome path be $\{\mathbf{a}^t(\boldsymbol{\sigma}(t))\}_{t=1}^\infty$, wherein a strategy $\sigma_i(t)$ at

stage t is chosen from S_i based on the history of the game at $t - 1$, *i.e.* $\sigma_i(1) \in S_i$ and for $t > 1$, $\sigma_i(t) : (S_M \times S_P)^{t-1} \rightarrow S_i$. This characterizes the player's choice of action $a_i^t(\sigma(t))$ at stage t as a function of the information gathered from the previous $t - 1$ actions. This history is public and is known to both players as each can observe the other's action directly at every stage.

Now, suppose the cooperative payoff V_c cannot be supported by any equilibrium in the repeated game, such that M will always find it profitable to cheat in the first stage even if he has to bear the subsequent (minimaxing) punishment forever. That is:

$$V_c + V_c\delta_M + V_c\delta_M^2 + \dots < \bar{V} + \hat{V}\delta_M + \hat{V}\delta_M^2 + \dots$$

$$\Leftrightarrow V_c < \bar{V}(1 - \delta_M) + \hat{V}\delta_M$$

$$\Leftrightarrow \delta_M < \tilde{\delta} = \frac{\bar{V} - V_c}{\bar{V} - \hat{V}}$$

Knowing that M cannot anymore be trusted to cooperate since $\delta_M < \tilde{\delta}$, P on her part will simply minimax M right from the start, inducing him to get only \hat{V} . And if P does not minimax M , she herself will be minimaxed by M by virtue of M 's impatience, and so on. Notice that this “mutual” minimaxing is enough to reduce the game to the stage-game equilibrium (\hat{V}, \hat{V}) ⁶. Thus, from a sustainable Pareto optimal outcome (V_c, V_c) , the game reverts to the inferior pure-strategy equilibrium when M 's discount factor goes below $\tilde{\delta}$. If actions are made contingent on the result of some public randomization, some individually rational payoffs to M may still be sustained in equilibrium despite his low discount factor (*e.g.* those (V_M, V_P) where $V_M > V_c$ and $V_P \geq \hat{V}$). However, most of them fail

⁶In some cases, a binding minimaxing punishment scheme (\hat{V}, \hat{V}) demands that a continuously increasing penalty at every succeeding stage is established in order that punishment is surely inflicted to avoid being punished more severely in the next stage. This complication on higher-order punishments was resolved by Abreu (1988) by introducing a simple punishment strategy that does not depend on the previous sequence of deviations and which can be supported in perfect equilibrium. Furthermore, such minimaxing actions may require mixed strategies in general and one has to assume that they are observable to obtain the Folk Theorem result. However this assumption is not indispensable as argued by Fudenberg and Maskin (1991) since the same result can also be achieved by employing over time a cyclical set of alternating pure actions with the appropriate frequency.

particularly those $V_M \in [\hat{V}, V_c]$ since none of these payoffs can in any way deter M from deviating in the first period.

4. Tolerant Strategies

It is expected from the Folk Theorem that some individually rational payoffs in the stage-game cannot anymore be sustained in equilibrium in repeated games when there exists an impatient player. This loss of equilibria is explained by the fact that both players continue to hold on to a trigger strategy that aims for the optimal cooperative outcome when it is no longer attainable. Consequently, the strategy ceases to be efficient since it automatically leads the game towards its punishment path that immediately penalizes both players and only eliminates the possibility of extracting some other feasible gains.

This scenario, however, is changed when the patient player P (with $\delta_P > \tilde{\delta}^7$) abandons the original strategy and concedes to adopt a tolerant trigger strategy. Although this may provide unequal and suboptimal yields (for P) in general, the generation of Pareto-superior equilibria is shown to be a worthwhile consolation as this can even approximate the cooperative outcome. Formally, a tolerant trigger strategy is defined as follows:

Definition 1. A *tolerant trigger strategy* (TTS) is an action profile $\{\mathbf{a}^t(\boldsymbol{\sigma}(t))\}_{t=1}^\infty$ in a repeated game $\Gamma^\infty(\delta_M, \delta_P)$ which satisfies the following conditions:

(i) there exists a certain strategy $(\bar{\sigma}_M(t), \underline{\sigma}_P(t)) \in S_M \times S_P$ that generates stage payoffs $\mu_M(\mathbf{a}^{t'}) > V_c$ and $\mu_P(\mathbf{a}^{t'}) < \hat{V}$ at some stages $t' \in \{1, 2, 3, \dots\}$, where $\mathbf{a}^{t'} := \mathbf{a}^{t'}(\bar{\sigma}_M(t'), \underline{\sigma}_P(t'))$.

(ii) $(1 - \delta_P) \sum_{t=1}^\infty \delta_P^{t-1} \mu_P(\mathbf{a}^t) > \hat{V}$.

(iii) once a deviation occurs at any time d , a minimaxing punishment (\hat{V}, \hat{V}) is played from time $d + 1$ onwards.

The first condition requires the existence of some stage-payoffs that are lower than the individually rational level (which allows the other player to earn higher

⁷We simply apply here a strict rather than weak inequality for the purpose of simplifying our results.

than the cooperative yields) while the second guarantees that the average discounted payoff of the tolerant player over the entire repeated game is above the individually rational level. The third is the typical trigger punishment path.

I characterize every TTS profile $\{a_i^t(\sigma(t))\}_{t=1}^\infty$ as a combination of *contract regime* which is the phase when both players continue to play the game according to what they have initially agreed on and a *punishment regime* that immediately sets in after a breach from the contract regime or from the same punishment regime (as will be discussed in section 7) has occurred. At this stage, it would be convenient to focus our analysis on infinitely repeated prisoner's dilemma game $\Gamma_{pd}^\infty(\delta_M, \delta_P)$ whose punishment regime is stable, being always a Nash equilibrium. This saves us from worries about the credibility of punishment and allows us to put more attention on the inherent difficulty that the contract regime of a TTS brings. One can see that unlike in the normal trigger strategy, the contract or the initial path of TTS no longer constitutes of playing the same action throughout its phase and can even take many different forms. Consequently, the continuation strategies at every subgame can differ since payoffs within the contract regime are not anymore the same. A simple classification of TTS profiles is presented below where we define payoffs during tolerant stages as $\mu_M(\mathbf{a}^t) = \bar{V}$ and $\mu_P(\mathbf{a}^t) = \underline{V}$, where again $\underline{V} < \hat{V} < V_c < \bar{V}$.

Definition 2.

(a) A *periodic tolerant trigger strategy* $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$ (or PTTS) is a TTS profile that has a contract regime of playing alternately k stages of tolerance with stage payoff $\mu(\mathbf{a}^t) = (\bar{V}, \underline{V})$ and then j stages of cooperation with stage payoff (V_c, V_c) over the game $\Gamma_{pd}^\infty(\delta_M, \delta_P)$. We write $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$ to denote a PTTS that starts with cooperative stages.

(b) A *non-periodic tolerant trigger strategy* is a TTS profile that starts with either k stages of tolerance followed by infinite stages of cooperation or with j stages of cooperation followed by tolerant stages thereafter.

As our analysis is confined only on discrete time between stages, we shall set k and j to be finite elements of the set of positive integers, \mathbb{Z}^+ .

Proposition 1. *For any $\delta_M < \tilde{\delta}$ and $\delta_P < 1$ in $\Gamma_{pd}^\infty(\delta_M, \delta_P)$, it is impossible to sustain a non-periodic tolerant trigger strategy in equilibrium.*

Proof:

Suppose it is possible. Then, in any of the following two cases, there exists a scenario when both players prefer to stick to the non-periodic TTS than to deviate from it.

Case A: (Tolerance before cooperation)

Examine M 's behavior. Notice that if M were to deviate, it has to be in the stage of cooperation since deviating when he is tolerated will only give him a lower payoff (i.e. $V_c < \bar{V}$). Thus, for M to remain faithful to the strategy, his payoff must be at least as much as the payoff he gets when he deviates at any cooperative stage.

$$(1-\delta_M) \left(\sum_{t=1}^k \bar{V} \delta_M^{t-1} + \sum_{t=k+1}^{\infty} V_c \delta_M^{t-1} \right) \geq (1-\delta_M) \left(\frac{\sum_{t=1}^k \bar{V} \delta_M^{t-1} + \sum_{t=k+1}^{k+q} V_c \delta_M^{t-1}}{\bar{V} \delta_M^{k+q} + \sum_{t=k+q+2}^{\infty} \hat{V} \delta_M^{t-1}} \right),$$

for all $q \in \{0, 1, 2, \dots\}$ and $k \in \mathbb{Z}^+$, where q is the number of stages of cooperation just before defecting. This implies now the following:

$$\begin{aligned} \Rightarrow V_c \delta_M^k &\geq V_c \delta_M^k (1 - \delta_M^q) + \bar{V} \delta_M^{k+q} (1 - \delta_M) + \hat{V} \delta_M^{k+q+1} \\ \Rightarrow \delta_M^{k+q+1} (\bar{V} - \hat{V}) &\geq \delta_M^{k+q} (\bar{V} - V_c) \\ \Rightarrow \delta_M &\geq \frac{\bar{V} - V_c}{\bar{V} - \hat{V}} = \tilde{\delta}, \text{ a contradiction.} \end{aligned}$$

Case B: (Cooperation before tolerance)

Examine P 's behavior. For P to stick to the (non-periodic) tolerant strategy, the payoff must be at least as much as the payoff she gets in any possible stage of deviation. Consider the possible deviation at the stage when P is about to start tolerating M (i.e. at $t = j + 1$ and that no deviation has occurred in the past).

We see that the condition not to deviate at this stage, *i.e.*

$$(1 - \delta_P) \left(\sum_{t=1}^j V_c \delta_P^{t-1} + \sum_{t=j+1}^{\infty} \underline{V} \delta_P^{t-1} \right) \geq (1 - \delta_P) \left(\sum_{t=1}^j V_c \delta_P^{t-1} + \sum_{t=j+1}^{\infty} \hat{V} \delta_P^{t-1} \right)$$

cannot hold since $\sum_{t=j+1}^{\infty} \underline{V} \delta_P^{t-1} < \sum_{t=j+1}^{\infty} \hat{V} \delta_P^{t-1}$, for all $j \in \mathbb{Z}^+$ and $\delta_P < 1$.

■

Remark. The proof in case A is general since it considers all of M 's possible deviation in any of the cooperative stages, whereas case B picks up only a stage where P 's defection is imminent. In both cases, it is shown that non-periodic TTS breaks down within a given player, independent of the other player's capacity to hold on to the strategy.

In the succeeding subsections, it is presented that perfect equilibrium can be generated under PTTS.

4.1. Tolerance before cooperation

Definition 3. In a game $\Gamma_{pd}^{\infty}(\delta_M, \delta_P)$, any PTTS $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$ is supported by a *subgame perfect equilibrium*, if for any strategy $\sigma'_i(d) \in S_i$ that differs from strategy $\sigma_i^{\langle k,j \rangle}$ at time d onwards, for all $k, j \in \mathbb{Z}^+$ and $i \in \{M, P\}$, we have:

$$\begin{aligned} (1 - \delta_i) \sum_{t=1}^{\infty} \delta_i^{t-1} \mu_i(\mathbf{a}^t(\sigma^{\langle k,j \rangle})) &\geq (1 - \delta_i) \sum_{t=1}^d \delta_i^{t-1} \mu_i(\mathbf{a}^t(\sigma'_i(d), \sigma_{-i}^{\langle k,j \rangle})) \\ &\quad + (1 - \delta_i) \sum_{t=d+1}^{\infty} \delta_i^{t-1} \mu_i(\varphi_i, \varphi_{-i}), \end{aligned}$$

where $(\varphi_i, \varphi_{-i})$ is the action vector of minimaxing punishment.

This definition of subgame perfection suffices to hold for prisoner's dilemma since its minimaxing punishment path is always Nash equilibrium. Thus, there is indeed no incentive for players to deviate during the punishment regime at any subgame. This allows us now with the task of ensuring only that deviation from the contract path at any stage is never profitable.

However, complexity still arises since continuation payoffs at any time d vary over the infinite period and most subgames within the contract regime are no longer identical to the original game. Apart from this, the stage payoffs of the two players are non-symmetric which impels us to deal with each one's payoff incentives separately before pinning down the set of perfect equilibrium points.

When the PTTS $(\sigma_M^{(k,j)}, \sigma_P^{(k,j)})$ is followed consistently over the entire game, the respective average discounted payoff to M and P are:

$$\Pi_M^{(k,j)} = (1 - \delta_M) \sum_{T=0}^{\infty} \left(\sum_{t=1}^k \bar{V} \delta_M^{t-1} + \sum_{t=k+1}^{k+j} V_c \delta_M^{t-1} \right) \delta_M^{T(k+j)} = \frac{(\bar{V} - V_c) (1 - \delta_M^k)}{1 - \delta_M^{k+j}} + V_c \quad (1)$$

and

$$\Pi_P^{(k,j)} = (1 - \delta_P) \sum_{T=0}^{\infty} \left(\sum_{t=1}^k \underline{V} \delta_P^{t-1} + \sum_{t=k+1}^{k+j} V_c \delta_P^{t-1} \right) \delta_P^{T(k+j)} = \frac{(\underline{V} - V_c) (1 - \delta_P^k)}{1 - \delta_P^{k+j}} + V_c \quad (2)$$

In order for the strategy $(\sigma_M^{(k,j)}, \sigma_P^{(k,j)})$ to be sustained in the game $\Gamma_{pd}^{\infty}(\delta_M, \delta_P)$, $\Pi_i^{(k,j)}$ must be at least as much as the average discounted payoff of i over the entire game when he/she decides to deviate at some time d ⁸. Although this requires us to identify the condition for the possible deviation at each and every stage of the infinite game, the following lemmas (1-3) allow us to simplify our investigation. The first limits our investigation from infinite number of stages into just the first $k+j$ stages. The second asserts that deviation cannot occur during tolerant stages (k stages) while the third shows the monotonic property of payoffs when deviating during the cooperative stages (j stages).

Lemma 1.

The condition not to deviate at the n^{th} stage of a PTTS, where n is an integer from 1 to $k+j$, is the same condition that holds for any $n + T(k+j)^{th}$ stage, where T is any positive integer.

⁸This method of comparing the entire-game yield between the no-deviation and the d th period-deviation case should provide same result as when comparing only their continuation payoffs from d , since their average discounted payoff before d are the same. I refrain from the typical use of continuation payoffs for computational simplicity.

Proof:

Let $\langle x_{(s)} \rangle_{s=1}^{k+j}$ be an arrangement of payoffs for the first $k+j$ stages with a discounted sum of $S_{\langle k,j \rangle} = x_{(1)} + x_{(2)}\delta + \dots + x_{(k+j)}\delta^{k+j-1}$. When no deviation occurs from the periodic tolerant strategy, $\langle x_{(s)} \rangle_{s=1}^{k+j}$ is repeated infinitely times and has a discounted sum of $S_{\langle k,j \rangle} + S_{\langle k,j \rangle}\delta^{k+j} + S_{\langle k,j \rangle}\delta^{2(k+j)} + \dots$. A deviation at n^{th} stage, where $n \leq k+j$, has a payoff profile of $\langle \langle x_{(s)} \rangle_{s=1}^{n-1}, \bar{x}_{(n)}, \langle \hat{x}_{(s)} \rangle_{s=n+1}^{\infty} \rangle$, where $\bar{x}_{(n)}$ is the payoff from deviating at n and \hat{x} is the subsequent punishment payoff the deviant receives. Denoting the discounted sum of this deviation path as S_D , we write the condition for sticking to the strategy at stage n as $S_{\langle k,j \rangle} + S_{\langle k,j \rangle}\delta^{k+j} + S_{\langle k,j \rangle}\delta^{2(k+j)} + \dots \geq S_D$.

Now, observe that when deviation occurs at $n + (k+j)^{th}$ stage, the discounted sum of the deviation path is $S_{\langle k,j \rangle} + S_D\delta^{k+j}$; while at $n + 2(k+j)^{th}$ stage, it is $S_{\langle k,j \rangle} + S_{\langle k,j \rangle}\delta^{k+j} + S_D\delta^{2(k+j)}$; and so on. Thus, the condition for not deviating at $n + T(k+j)^{th}$ stage, for any positive integer T , is as follows:

$$S_{\langle k,j \rangle} + S_{\langle k,j \rangle}\delta^{k+j} + S_{\langle k,j \rangle}\delta^{2(k+j)} + \dots \geq \underbrace{S_{\langle k,j \rangle} + S_{\langle k,j \rangle}\delta^{k+j} + \dots + S_{\langle k,j \rangle}\delta^{(T-1)(k+j)}}_{T \text{ number of terms}} + S_D\delta^{T(k+j)}.$$

Cancelling the first T terms on both sides, we get:

$$S_{\langle k,j \rangle}\delta^{T(k+j)} + S_{\langle k,j \rangle}\delta^{(T+1)(k+j)} + S_{\langle k,j \rangle}\delta^{(T+2)(k+j)} + \dots \geq S_D\delta^{T(k+j)}.$$

Then, by dividing both sides by $\delta^{T(k+j)}$, we obtain the same condition $S_{\langle k,j \rangle} + S_{\langle k,j \rangle}\delta^{k+j} + S_{\langle k,j \rangle}\delta^{2(k+j)} + \dots \geq S_D$. ■

Lemma 2.

Both players will not find it profitable under PTTS to deviate during any stage of tolerance.

Proof:

By definition, player P 's average payoff in sticking to the strategy is higher than the minimax level, \hat{V} . Clearly, to deviate during any of the prescribed tolerant stages will give her an average payoff of at most \hat{V} , that is, $(1 - \delta_P)(\sum_{t=1}^{k'} \underline{V}\delta_P^{t-1} + \sum_{t=k'+1}^{\infty} \hat{V}\delta_P^{t-1}) \leq \hat{V}$, where $k' \leq k$ is the number of tolerant stages conceded before deviating in the next stage. If $k' = 0$, then the game reverts to the minimax

equilibrium where P gets exactly \hat{V} . For player M , to deviate at the stage when he is tolerated only gives him a lower payoff $V_c < \bar{V}$. Moreover, the fact that his future stage payoffs are reverted to the minimax level after such deviation only deprives him of getting higher average income. ■

The moment P deviates during one of these tolerant stages, she loses the possibility of getting the cooperation of M in the future which could give her higher payoff, enough to even cover her losses during those tolerant stages. Similarly, M would not think of deviating during periods of tolerance since he is being tolerated to get high returns. Hence, we are left with the cooperative stages as the only possible periods where deviation can occur. In particular, we look for the highest payoff one can derive from all those possible deviations during the cooperative stages. This is presented formally as follows:

$$D_M^{\langle k,j \rangle} = (1 - \delta_M) \sum_{t=1}^k \bar{V} \delta_M^{t-1} + \max_{q \in \{0,1,\dots,j-1\}} \theta_M \left(V_c(q), \bar{V}, \hat{V}, \delta_M \right), \quad (3)$$

$$\text{where } \theta_M(\cdot) = (1 - \delta_M) \left(\sum_{t=k+1}^{k+q} V_c \delta_M^{t-1} + \bar{V} \delta_M^{k+q} + \sum_{t=k+q+2}^{\infty} \hat{V} \delta_M^{t-1} \right)$$

$$D_P^{\langle k,j \rangle} = (1 - \delta_P) \sum_{t=1}^k \bar{V} \delta_P^{t-1} + \max_{r \in \{0,1,\dots,j-1\}} \theta_P \left(V_c(r), \bar{V}, \hat{V}, \delta_P \right), \quad (4)$$

$$\text{where } \theta_P(\cdot) = (1 - \delta_P) \left(\sum_{t=k+1}^{k+r} V_c \delta_P^{t-1} + \bar{V} \delta_P^{k+r} + \sum_{t=k+r+2}^{\infty} \hat{V} \delta_P^{t-1} \right).$$

The function $\theta_i(\cdot)$ depicts the average discounted payoff from deviating during the cooperative stages while the imbedded parameters q and r are the players' respective number of stages given to cooperation just before deviating from the strategy. Note that when q and r are equal to j , this means that deviation occurs at the stage of tolerance which was already ruled out in Lemma 2. Lemma 3 allows us to determine the maximum entire-game payoff one can obtain from deviating at any time during these cooperative stages.

Lemma 3. (*Monotonicity*)

- (i) $\theta_M(\cdot)$ is monotone decreasing in q .
- (ii) $\theta_P(\cdot)$ is monotone increasing in r .

Proof:

$$\begin{aligned}
\text{(i)} \quad \theta_M(\cdot) &= (1 - \delta_M) \left(\sum_{t=k+1}^{k+q} V_c \delta_M^{t-1} + \bar{V} \delta_M^{k+q} + \sum_{t=k+q+2}^{\infty} \hat{V} \delta_M^{t-1} \right) \\
&= V_c \delta_M^k (1 - \delta_M^q) + \bar{V} \delta_M^{k+q} (1 - \delta_M) + \hat{V} \delta_M^{k+q+1} \\
&= V_c \delta_M^k + (\bar{V} - V_c) \delta_M^{k+q} - (\bar{V} - \hat{V}) \delta_M^{k+q+1} \\
&= V_c \delta_M^k + \delta_M^{k+q} \left[(\bar{V} - V_c) - (\bar{V} - \hat{V}) \delta_M \right] \\
&= V_c \delta_M^k + \delta_M^{k+q} (\bar{V} - \hat{V}) (\tilde{\delta} - \delta_M).
\end{aligned}$$

Both terms in the last equation are positive. And since $0 < \delta_M < 1$, δ_M^{k+q} decreases in q and so as $\theta_M(\cdot)$.

$$\begin{aligned}
\text{(ii)} \quad \theta_P(\cdot) &= (1 - \delta_P) \left(\sum_{t=k+1}^{k+r} V_c \delta_P^{t-1} + \bar{V} \delta_P^{k+r} + \sum_{t=k+r+2}^{\infty} \hat{V} \delta_P^{t-1} \right) \\
&= V_c \delta_P^k (1 - \delta_P^r) + \bar{V} \delta_P^{k+r} (1 - \delta_P) + \hat{V} \delta_P^{k+r+1} \\
&= V_c \delta_P^k + \delta_P^{k+r} \left[(\bar{V} - V_c) - (\bar{V} - \hat{V}) \delta_P \right] \\
&= V_c \delta_P^k + \delta_P^{k+r} (\bar{V} - \hat{V}) (\tilde{\delta} - \delta_P)
\end{aligned}$$

Given that $1 > \delta_P > \tilde{\delta}$, the last term is always negative and therefore any increase in r reduces the negative value of the last term which increases $\theta_P(\cdot)$. ■

By Lemma 3, we obtain the highest values of $\theta_M(\cdot)$ and $\theta_P(\cdot)$ when $q = 0$ and $r = j - 1$, respectively, hence:

$$D_M^{\langle k, j \rangle} = \bar{V} - (\bar{V} - \hat{V}) \delta_M^{k+1} \quad \text{and} \quad (5)$$

$$D_P^{\langle k, j \rangle} = \underline{V} - (\underline{V} - V_c) \delta_P^k + (\bar{V} - V_c) \delta_P^{k+j-1} - (\bar{V} - \hat{V}) \delta_P^{k+j}. \quad (6)$$

Thus, the no-deviation condition for the strategy $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$, for any $k, j \in \mathbb{Z}^+$, is characterized by the inequality $\Pi_i^{\langle k,j \rangle} \geq D_i^{\langle k,j \rangle}$, for both i . Consequently, this condition provides a range of values of δ_M and δ_P that can support the fidelity of players to a periodic tolerant contract parameterized by k and j . Caution however should be made since some outcomes induced by these periodic contracts may even fail to be individually rational.

4.2. Cooperation before tolerance

An impatient player can also be made to cooperate initially despite having $\delta_M < \tilde{\delta}$, provided that the contract ensures that he be tolerated afterwards, in a periodic fashion *i.e.* $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$. An immediate question that can arise is how different is this strategy from the previously discussed $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$ in characterizing the set of no-deviation outcomes. One can observe immediately that their payoff yields are different in the sense that when a PTTS $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$ is followed faithfully over the entire game, the respective average discounted payoff to M and P are:

$$\Pi_M^{\langle j,k \rangle} = (1 - \delta_M) \sum_{T=0}^{\infty} \left(\sum_{t=1}^j V_c \delta_M^{t-1} + \sum_{t=j+1}^{j+k} \bar{V} \delta_M^{t-1} \right) \delta_M^{T(j+k)} = \frac{(V_c - \bar{V})(1 - \delta_M^j)}{1 - \delta_M^{j+k}} + \bar{V} \quad (7)$$

and

$$\Pi_P^{\langle j,k \rangle} = (1 - \delta_P) \sum_{T=0}^{\infty} \left(\sum_{t=1}^j V_c \delta_P^{t-1} + \sum_{t=j+1}^{j+k} \underline{V} \delta_P^{t-1} \right) \delta_P^{T(j+k)} = \frac{(V_c - \underline{V})(1 - \delta_P^j)}{1 - \delta_P^{j+k}} + \underline{V} \quad (8)$$

Notice that these results are different from the values of $\Pi_M^{\langle k,j \rangle}$ and $\Pi_P^{\langle k,j \rangle}$, as seen from (1) and (2). Interestingly however, the conditions that allow the strategy $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$ to generate no-deviation outcomes are the same with the strategy $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$. In brief, we say that $\Pi_i^{\langle k,j \rangle} \geq D_i^{\langle k,j \rangle}$ and $\Pi_i^{\langle j,k \rangle} \geq D_i^{\langle j,k \rangle}$ are identical, as shown in the following proposition.

Proposition 2. (*Equivalence*) For any $k, j \in \mathbb{Z}^+$, the PTTS $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$

and $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$ constitute the same range of values of δ_M and δ_P that can support the no-deviation condition during the contract regime of the repeated game. These values are defined by the following conditions:

$$\text{For player } M: \quad \delta_M^j \geq \frac{\tilde{\delta} - \delta_M}{\tilde{\delta} - \delta_M^{k+1}} \quad (9)$$

$$\text{For player } P: \quad \delta_P^j \leq \delta_P^{-k} + A \frac{1 - \delta_P^k}{\delta_P^{k-1} (\delta_P - \tilde{\delta})}, \text{ where } A = \frac{V - V_c}{\bar{V} - \hat{V}} < 0 \quad (10)$$

Proof: Appendix

From the results of Lemmas 1-3 and Proposition 2, the characterization of the set of perfect equilibrium outcomes can now be expressed in the following theorem.

Theorem 1. *In a game $\Gamma_{pd}^\infty(\delta_M, \delta_P)$, where $\delta_M < \tilde{\delta} < \delta_P$ and where $\tilde{\delta}$ is the minimum level of discount factor that can support a cooperative outcome, there exists a (subgame) perfect equilibrium characterized by PTTS $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$ and $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$, where $k, j \in \mathbb{Z}^+$,*

(a) *for all $\delta_M \in (\underline{\delta}_M, \tilde{\delta})$ and $\delta_P \in (\underline{\delta}_P, 1)$, where $\underline{\delta}_M \in (\frac{\tilde{\delta}}{1+\tilde{\delta}}, \tilde{\delta})$ and $\underline{\delta}_P \in (\tilde{\delta}, 1)$ and*

(b) *with average discounted payoffs of $\Pi_M^{\langle k,j \rangle}, \Pi_M^{\langle j,k \rangle} \in (V_c, \bar{V})$ and $\Pi_P^{\langle k,j \rangle}, \Pi_P^{\langle j,k \rangle} \in (\hat{V}, V_c)$.*

Clearly, by the assertion of Theorem 1(b), the classic Folk Theorem result is not obtained here since payoffs between \hat{V} and V_c are not feasible to M . Nonetheless, for those payoffs restored in perfect equilibrium, the theorem depicts well the range of discount factors that can support them.

Proof:

(a)

(Step 1) Recall first that any deviation at any stage of a prisoner's dilemma game is responded by a minimaxing Nash punishment, making the punishment regime always binding. Thus, one only needs to guarantee that there will also be no incentive to deviate during the contract regime. Lemmas 1, 2, and 3 reduce

this condition of no-deviation to $\Pi_i^{(k,j)} \geq D_i^{(k,j)}$ while Proposition 2 shows that this is equivalent to $\Pi_i^{(j,k)} \geq D_i^{(j,k)}$ and is brought down to the equilibrium constraints for each player, as depicted in (9) and (10). To complete the characterization of perfect equilibrium payoffs, we invoke the definition of TTS, i.e. $\Pi_P^{(k,j)} > \hat{V}$ and $\Pi_P^{(j,k)} > \hat{V}$ (individually rational condition (IRC)). We show later in the proof of Theorem 1 (b) that these payoffs above \hat{V} , while fulfilling condition (10), do certainly exist.

Our goal in the next step is to pin down the lowest possible values of δ_M and δ_P on which perfect equilibrium can still be satisfied. A key to this is the result of Lemma 4, presented at the end of the proof.

(Step 2.1) Set a correspondence $\Phi_M : \mathbb{Z}^+ \times \mathbb{Z}^+ \rightarrow \varsigma \left((0, \tilde{\delta}) \right)$ defined by⁹

$$\Phi_M(k, j) = \left\{ \left(\delta_M, \tilde{\delta} \right) \subseteq (0, \tilde{\delta}) \left| \begin{array}{l} \delta_M^j \geq \frac{\tilde{\delta} - \delta_M}{\tilde{\delta} - \delta_M^{k+1}}, \text{ for a given } (k, j) \in \mathbb{Z}^+ \times \mathbb{Z}^+ \\ \text{and } \tilde{\delta} \in (0, 1) \end{array} \right. \right\}.$$

Note that the lowest δ_M of the interval $(\delta_M, \tilde{\delta})$ is solved by the equality part of (9).

(i) Fix j at j_o . Then, as k increases, δ_M decreases (by Lemma 4(i)), which expands the set $\Phi_M(k, j_o)$. Thus, Φ_M is monotone increasing in k , i.e. $\Phi_M(k, j_o) \subset \Phi_M(k+1, j_o)$.

(ii) Fix k at k_o . Then, as j increases, δ_M increases (by Lemma 4(i)) and approaches $\tilde{\delta}$. Thus, Φ_M is monotone decreasing in j , i.e. $\Phi_M(k_o, j) \supset \Phi_M(k_o, j+1)$.

From (i) and (ii), Φ_M is largest when $k \rightarrow \infty$ and $j = 1$; hence, we solve from (9) that the least δ_M , i.e. $\underline{\delta}_M$, is $\frac{\tilde{\delta}}{1+\tilde{\delta}}$. Thus, for any finite $k, j \in \mathbb{Z}^+$ that satisfies (9), $\delta_M \in \left(\frac{\tilde{\delta}}{1+\tilde{\delta}}, \tilde{\delta} \right)$.

(iii) Finally, we show that P can likewise admit a profile where $j = 1$ and $k \rightarrow \infty$ by satisfying (10) and the IRC. Consider the strategy profile $\sigma^{(1,k)}$. Then, $k \rightarrow \infty \Rightarrow \Pi_P^{(1,k)} \geq D_P^{(1,k)} > \hat{V}$, making (10) the only binding constraint. We write

⁹ $\varsigma \left((0, \tilde{\delta}) \right)$ reads as the power set of the interval $(0, \tilde{\delta})$.

(10) as $\delta_P^{k+j} \leq 1 + A \frac{\delta_P(1-\delta_P^k)}{\delta_P - \tilde{\delta}}$ and as $k \rightarrow \infty$, this implies that $\sigma^{\langle 1, k \rangle}$ is supported for as long as $A \geq -1 + \frac{\tilde{\delta}}{\delta_P}$.

(Step 2.2) Similarly, we set a correspondence $\Phi_P : \mathbb{Z}^+ \times \mathbb{Z}^+ \rightarrow \varsigma\left(\left(\tilde{\delta}, 1\right)\right)$ defined by

$$\Phi_P(k, j) = \left\{ (\delta_P, 1) \subseteq \left(\tilde{\delta}, 1\right) \left| \begin{array}{l} \delta_P^j \leq \delta_P^{-k} + A \frac{1-\delta_P^k}{\delta_P^{k-1}(\delta_P - \tilde{\delta})}, \text{ where } (k, j) \in \mathbb{Z}^+ \times \mathbb{Z}^+, \\ \tilde{\delta} \in (0, 1) \text{ and } A = \frac{V - V_c}{\bar{V} - \bar{V}} \end{array} \right. \right\}.$$

We argue in a similar fashion as above where in this case the lowest δ_P of the interval $(\delta_P, 1)$ is solved by the equality condition of (10). By Lemma 4 (ii), it implies that δ_P increases in k , therefore $\Phi_P(k, j)$ is monotone decreasing in k *i.e.* $\Phi_P(k, j_o) \supset \Phi_P(k+1, j_o)$. On the other hand, $\Phi_P(k, j)$ is monotone increasing in j *i.e.* $\Phi_P(k_o, j) \subset \Phi_P(k_o, j+1)$ since δ_P decreases in j . Thus, set $\Phi_P(k, j)$ is largest when $k = 1$ and $j \rightarrow \infty$ (and this can easily pass the IRC, *e.g.* $\Pi_P^{\langle j, 1 \rangle} > \hat{V}$ for $j \rightarrow \infty$). By plugging in these values in the equality of (10), we obtain the lowest δ_P as $\underline{\delta}_P = \frac{(A+1) \pm \sqrt{(A+1)^2 - 4A\tilde{\delta}}}{2A}$. Since $\lim_{A \rightarrow -\infty} \underline{\delta}_P = 1 > \lim_{A \rightarrow -1} \underline{\delta}_P \sqrt{\tilde{\delta}} > \lim_{A \rightarrow 0} \underline{\delta}_P = \tilde{\delta}$, this implies that for any finite $k, j \in \mathbb{Z}^+$ that satisfies (10) and for $A < 0$, $\delta_P \in (\tilde{\delta}, 1)$. Finally, we show that strategy profiles $\sigma^{\langle j, 1 \rangle}$ and $\sigma^{\langle 1, j \rangle}$, where $j \rightarrow \infty$, are both admissible to player M . Suppose $\delta_M \rightarrow \tilde{\delta}$, then we see that $\delta_M^j \geq \frac{\tilde{\delta} - \delta_M}{\tilde{\delta} - \delta_M^2} \Leftrightarrow \tilde{\delta}^j \geq 0$ is satisfied even if $j \rightarrow \infty$.

(b)

Since $\Pi_M^{\langle k, j \rangle} = \frac{(\bar{V} - V_c)(1 - \delta_M^k)}{1 - \delta_M^{k+j}} + V_c$ and $\frac{1 - \delta_M^k}{1 - \delta_M^{k+j}} \in (0, 1)$ for any $k, j \in \mathbb{Z}^+$ and $\delta'_M \in (\underline{\delta}_M, \tilde{\delta})$, we obtain $\Pi_M^{\langle k, j \rangle} \in (V_c, \bar{V})$. A similar argument also allows us to show that $\Pi_M^{\langle j, k \rangle} \in (V_c, \bar{V})$.

For player P with $\delta'_P \in (\underline{\delta}_P, 1)$, the IRC is binding only for some pairs of $k, j \in \mathbb{Z}^+$, such that $\Pi_P^{\langle k, j \rangle} = \frac{(V - V_c)(1 - \delta_P^k)}{1 - \delta_P^{k+j}} + V_c > \hat{V} \Leftrightarrow k < \bar{k}(j) = \left(\frac{1}{\log \delta'_P}\right) \log \left(\frac{1 - \eta}{1 - \eta \delta_P^j}\right)$ where $\eta = \frac{V_c - \hat{V}}{V_c - \underline{V}} \in (0, 1)$. This shows that when $k < \bar{k}(j)$, we have $\Pi_P^{\langle k, j \rangle} > \hat{V}$ and since $\frac{1 - \delta_P^k}{1 - \delta_P^{k+j}} \in (0, 1)$, we clearly have $\Pi_P^{\langle k, j \rangle} \in (\hat{V}, V_c)$. Similarly, it can be shown that the IRC is satisfied for profile $\sigma_P^{\langle j, k \rangle}$ iff $j < \bar{j}(k) = \left(\frac{1}{\log \delta'_P}\right) \log \left(\frac{\eta}{1 - (1 - \eta) \delta_P^k}\right)$.

Thus, when $j < \bar{j}(k)$, we have $\Pi_P^{\langle j,k \rangle} \in (\hat{V}, V_c)$ since $\frac{1-\delta_P^j}{1-\delta_P^{j+k}} \in (0, 1)$. ■

Lemma 4. *The real roots δ_M and δ_P of the equations $\delta_M^j = \frac{\tilde{\delta}-\delta_M}{\tilde{\delta}-\delta_M^{k+1}}$ and $\delta_P^j = \delta_P^{-k} + A \frac{1-\delta_P^k}{\delta_P^{k-1}(\delta_P-\tilde{\delta})}$, respectively, that exist and belong to the interval $(0, 1)$, behave in the following manner with respect to k and j , for any $k, j \in \mathbb{Z}^+$:*

$$(i) \frac{\partial \delta_M}{\partial k} < 0 \text{ and } \frac{\partial \delta_M}{\partial j} > 0 \quad (ii) \frac{\partial \delta_P}{\partial k} > 0 \text{ and } \frac{\partial \delta_P}{\partial j} < 0$$

Proof: Appendix

4.3. Sets of perfect equilibrium payoffs

The result in Theorem 1 shows that for any discount factors between the interval $(\tilde{\delta}, \frac{\tilde{\delta}}{1+\tilde{\delta}})$ for M and between $(\tilde{\delta}, 1)$ for P , there exists a combination of finite number of stages of tolerance (k) and cooperation (j) that can generate perfect equilibrium payoffs. Any combination of k and j that satisfies Proposition 2 therefore generates a distinct set of possible equilibrium payoffs for both players under the strategies $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$ and $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$. Through the results of Lemma 5, we graph some of these sets in Figure 2. Note however that for some combinations of k and j , it is possible for P to generate payoffs lower than \hat{V} and yet admits the condition in Proposition 2. These strategy profiles that yield such payoffs violate the definition of TTS (individually rational condition) and are therefore not equilibrium outcomes. On the other hand, there is no danger for M to fail the individually rational condition since its payoff structure is always above V_c .

Lemma 5. *For any given $k, j \in \mathbb{Z}^+$, $\alpha \in \mathbb{Z}^+ \setminus \{1\}$, $\delta_M \in (\underline{\delta}_M, \tilde{\delta})$, and $\delta_P \in (\underline{\delta}_P, 1)$:*

$$\begin{aligned} (i) \quad & \Pi_M^{\langle k,j \rangle} > \Pi_M^{\langle j,k \rangle} \\ (ii) \quad & \Pi_P^{\langle j,k \rangle} > \Pi_P^{\langle k,j \rangle} \\ (iii) \quad & \Pi_M^{\langle j,k \rangle} > \Pi_M^{\langle \alpha j, \alpha k \rangle} \text{ and } \Pi_M^{\langle k,j \rangle} < \Pi_M^{\langle \alpha k, \alpha j \rangle} \\ (iv) \quad & \Pi_P^{\langle j,k \rangle} > \Pi_P^{\langle j, \alpha k \rangle} \text{ and } \Pi_P^{\langle k,j \rangle} > \Pi_P^{\langle \alpha k, j \rangle} \end{aligned}$$

Proof : Appendix

Remark. Lemma 5 illustrates well how the use of strategies $\sigma^{\langle k,j \rangle}$ and $\sigma^{\langle j,k \rangle}$ and the level of k and j affect the players average payoffs. The patient player,

for instance, obtains higher payoff under the strategy $\sigma^{\langle j,k \rangle}$ than in $\sigma^{\langle k,j \rangle}$ for any given k and j . Moreover, she is always better off when the number of tolerant stages (k) is kept as low as possible.

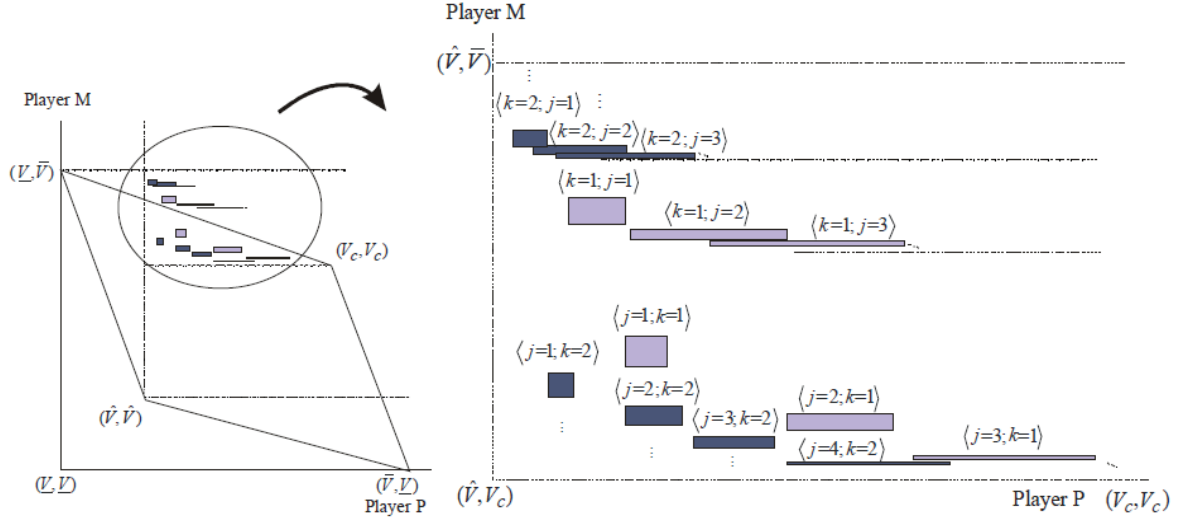


Figure 2: Each rectangular block in the figure corresponds to a set of perfect equilibrium payoffs generated by the strategy $\sigma^{\langle k,j \rangle}$ or $\sigma^{\langle j,k \rangle}$. The label $\langle j = 2; k = 1 \rangle$, for example, denotes that the strategy $\sigma^{\langle j=2,k=1 \rangle}$ is used.

5. Limit, Optimal, and Cooperative Tolerance

In this section, we study the case where a fixed δ'_M is set within the range $(\underline{\delta}_M, \tilde{\delta})$ vis-à-vis a δ_P that is very close to 1. The idea here is to answer the question how long can a very patient player tolerate a given impatient person in such a way that they still maintain an equilibrium payoff better than what they will receive in a single stage-game. This notion of limit tolerance explores the boundary to which PTTS can remain effective and enforceable. Moreover, it is also an interest to know how a patient player, in the course of setting offers of tolerance to the other, optimizes her returns. Thus, apart from generating superior equilibria, she is also concerned of maximizing her average income without making the other defect at any time of the game. Notice however that as the patient player tries to increase her payoff towards V_c , the other's payoff sinks towards V_c from

above. This, in the end, leads us to conjecture the attainability of a cooperative outcome.

Proposition 3. *For a given $\delta'_M \in (\underline{\delta}_M, \tilde{\delta})$ and any δ_P close to 1:*

(a) *(Limit Tolerance) the maximum level of tolerance P can render to M (regardless of payoffs) for any j number of cooperative stages is*

$$k^* = \overline{\mathbb{Z}} \left(j \frac{(V_c - \hat{V})}{(\hat{V} - \underline{V})} \right), \quad (11)$$

where $\overline{\mathbb{Z}}(x)$ is defined as the greatest integer less than x .

(b) *(Optimal Tolerance) player P maximizes her income from a PTTS by offering j^* stages of cooperation and a 1-stage tolerance, i.e. $\max \Pi_P = \Pi_P^{(j^*, 1)}$, where*

$$j^* = \tilde{\mathbb{Z}} \left(\frac{1}{\log \delta'_M} \log \left(\frac{\tilde{\delta} - \delta'_M}{\tilde{\delta} - \delta'^2_M} \right) \right), \quad (12)$$

and where $\tilde{\mathbb{Z}}(x)$ is the greatest integer less than or equal to x .

(c) *(Cooperative Tolerance) the cooperative outcome (V_c, V_c) can be approximated as $\delta'_M \rightarrow \tilde{\delta}$.*

Despite the asymmetric payoffs earned by players through a tolerant strategy profile, cooperative outcome can almost be reproduced under certain conditions (i.e. when δ'_M approaches $\tilde{\delta}$ and when j is set at a high level). This result is appreciated better when we recall that under the usual trigger strategy, when δ'_M becomes just below $\tilde{\delta}$, the effect is evident as the once achievable cooperative outcome can no longer be supported by pure strategies and the game immediately drops to a lower equilibrium (\hat{V}, \hat{V}) . Hence, in situations when $\delta'_M < \tilde{\delta}$, the PTTS not only can offer superior equilibria than the normal trigger strategy, but also can achieve an almost-cooperative outcome.

Proof:

(a) We are interested in finding

$$k^* = \sup \left\{ k \in \mathbb{Z}^+ \mid \left\{ \Pi_P^{(j,k)} \geq D_P^{(j,k)} \right\} \cap \left\{ \Pi_P^{(j,k)} > \hat{V} \right\}, \text{ for any } j \in \mathbb{Z}^+ \right\}$$

as we consider the profile $\sigma_P^{(j,k)}$. As $\sigma_P \rightarrow 1$, the condition $\left\{ \Pi_P^{(j,k)} \geq D_P^{(j,k)} \right\} \cap \left\{ \Pi_P^{(j,k)} > \hat{V} \right\}$ leads to the inequality requirement $\Pi_P^{(j,k)} = (V_c - \underline{V}) \frac{j}{j+k} + \underline{V} > \hat{V}$, for any $j, k \in \mathbb{Z}^+$. Rearranging, we obtain $k < j \left(\frac{V_c - \hat{V}}{\hat{V} - \underline{V}} \right)$. In the case of profile $\sigma_P^{(k,j)}$, the condition $\left\{ \Pi_P^{(k,j)} \geq D_P^{(k,j)} \right\} \cap \left\{ \Pi_P^{(k,j)} > \hat{V} \right\}$ leads to $\Pi_P^{(k,j)} = (\underline{V} - V_c) \frac{k}{k+j} + V_c > \hat{V}$ as $\sigma_P \rightarrow 1$ and for any $j, k \in \mathbb{Z}^+$. Rearranging, we get an identical result to profile $\sigma_P^{(j,k)}$ above. Thus, k^* is the highest integer less than $j \left(\frac{V_c - \hat{V}}{\hat{V} - \underline{V}} \right)$, for any $j \in \mathbb{Z}^+$. And since profiles $\sigma^{(j,k^*)}$ and $\sigma^{(k^*,j)}$ could easily pass the equilibrium requirements for M , we have completed the proof for (a).

(b) By Lemma 5(ii), P receives a higher payoff from the profile $\sigma^{(j,k)}$ than from $\sigma^{(k,j)}$. Moreover, for any j , $\Pi_P^{(j,k)}$ is highest when $k = 1$ by Lemma 5(iv). Now, we apply the condition $\left\{ \Pi_i^{(j,1)} \geq D_i^{(j,1)} \right\} \cap \left\{ \Pi_i^{(j,1)} > \hat{V} \right\}$ for both players. For P , we see from (a) that as $\delta_P \rightarrow 1$, this implies that $k < j \left(\frac{V_c - \hat{V}}{\hat{V} - \underline{V}} \right)$. Once this is satisfied, any further increase in the number of cooperative stages, j , would never induce P to deviate. This is not so, however, for M whose incentive not to deviate is given by $\delta_M^{j,k} \geq \frac{\tilde{\delta} - \delta'_M}{\tilde{\delta} - \delta'^{k+1}_M} \Leftrightarrow j \leq \left(\frac{1}{\log \delta'_M} \log \left(\frac{\tilde{\delta} - \delta'_M}{\tilde{\delta} - \delta'^{k+1}_M} \right) \right)$. Thus, for $k = 1$, the highest j that could still make M abide is $j^* = \tilde{\mathbb{Z}} \left(\frac{1}{\log \delta'_M} \log \left(\frac{\tilde{\delta} - \delta'_M}{\tilde{\delta} - \delta'^2_M} \right) \right)$, where $\tilde{\mathbb{Z}}(x)$ is the greatest integer less than or equal to x .

(c) Since P is never constrained by any increase of the number of cooperative stages for as long as $k < j \left(\frac{V_c - \hat{V}}{\hat{V} - \underline{V}} \right)$, then we see that $\lim_{\substack{\delta_P \rightarrow 1 \\ j \rightarrow \infty}} \Pi_P^{(j,k)} = \frac{(V_c - \underline{V})(1 - \delta_P^j)}{1 - \delta_P^{j+k}} + \underline{V} = V_c$. On the other end, one can also observe that $\lim_{j \rightarrow \infty} \Pi_M^{(j,k)} = V_c$. However, M has to satisfy the condition $\delta_M^{j,k} \geq \frac{\tilde{\delta} - \delta'_M}{\tilde{\delta} - \delta'^{k+1}_M}$ to generate perfect equilibrium points. Observe that even if $j \rightarrow \infty$, this condition for M can still be satisfied when $\delta'_M \rightarrow \tilde{\delta}$ since $\lim_{\delta'_M \rightarrow \tilde{\delta}} \frac{\tilde{\delta} - \delta'_M}{\tilde{\delta} - \delta'^{k+1}_M} = 0$, for any $k \in \mathbb{Z}^+$. ■

6. Equilibrium Restoration

As in the previous discussion, $\tilde{\delta}$ is defined as the smallest value of δ_M that can support the cooperative outcome V_c in equilibrium under the normal trigger strategy. In fact, as long as $\delta_M > \tilde{\delta}$, not only V_c but all V_M s in R that are above V_c can be sustained in equilibrium when there is public randomization. Notice also that each of these V_M s has its corresponding threshold discount factor that decreases as V_M increases. Now, we define:

$$\underline{\delta} = \inf \left\{ \tilde{\delta} \in (0, 1) \mid \tilde{\delta} = \frac{\bar{V} - V_M}{\bar{V} - \hat{V}}, \text{ for all } V_M \in \mathbb{R} \text{ such that vector } (V_M, \cdot) \in R \right\} \quad (13)$$

Therefore, when M is extremely impatient such that $\delta_M < \underline{\delta}$, there is no more payoff V_M in R (i.e. feasible and individually rational) that can achieve equilibrium. This complete loss of equilibrium payoffs can nonetheless be restored using PTTS which only manifests its greater efficiency over the normal trigger strategy in situations when an impatient player exists.

Theorem 2. (*Equilibrium Restoration*) *For some $\delta_M < \underline{\delta}$ such that there is no more individually rational payoff that can be sustained in equilibrium by the normal trigger strategy, there still exist some individually rational equilibria using PTTS.*

Proof:

Let $V_M^* \in R$ be the highest average payoff to M that can be sustained in perfect equilibrium using the normal trigger strategy with public randomization, if needed. Denote its corresponding discount factor threshold as $\tilde{\delta}$, such that for every $\delta_M = \tilde{\delta} - \varepsilon$, for small $\varepsilon > 0$, there is no more V_M in R that can be supported in equilibrium by the normal strategy. Pick two pure strategy payoffs: \bar{V} being the highest possible and V_c being any pure strategy payoff in R . Consider V_M' as an average payoff generated by a PTTS $\sigma^{(j,k)}$, such that $V_M' = \Pi_M^{(j,k)}(V_c, \bar{V}, \delta_M, k, j)$.¹⁰ By Theorem 1(a), we see however that equilibrium for M can still be obtained for all $\varepsilon \leq \tilde{\delta}^2 / (1 + \tilde{\delta})$. By Theorem 1(b), the

¹⁰By Proposition 2, the same result applies for both players if $\sigma^{(k,j)}$ instead of $\sigma^{(j,k)}$ is used.

individually rational condition is satisfied for both players. Thus, P receives simultaneously a payoff $V'_P = \Pi_P^{(j,k)}(V_c, \underline{V}, \delta_P, k, j)$ greater than \hat{V} , for some $k, j \in \mathbb{Z}^+$. And since payoff function is continuous in \mathbb{R}^2 , the vector (V'_M, V'_P) always exists for any $\tilde{\delta} - \varepsilon \leq \delta_M < \tilde{\delta}$ and $\delta_P < 1$. ■

7. Generalization to any Two-player Game

The main result in this section is presented in the following theorem.

Theorem 3. *The results in Theorem 1 continue to hold for any two-person game $\Gamma^\infty(\delta_M, \delta_P)$.*

Theorem 1 restricts the result to prisoner's dilemma game where punishment regime is intrinsically a Nash equilibrium while Theorem 3 generalizes this result to any two-person infinitely repeated game. The main feature of the proof of the latter is the typical *simple* punishment strategy proposed by Abreu (1988) that imposes the same punishment for any deviation and which does not lead to an escalating *hierarchy* of punishments as a result of dependence on past deviations. Fudenberg and Maskin (1986) used this method in a form of limited punishment which the proof of the above theorem tries also to employ.

The use of the minimaxing payoff (\hat{V}, \hat{V}) in prisoner's dilemma simplifies significantly the generation of equilibrium. In general, however, employing minimaxing payoff during the punishment regime may require mixed or correlated strategies since direct pure-strategy actions may not be possible. In this scenario, we simply assume that mixed strategies are observable or that there exists a public randomization device that can attain the minimaxing payoff (\hat{V}_M, \hat{V}_P) so that any deviation from these strategies can be detected. Unfortunately, the result of Fudenberg and Maskin (1991), which shows the possibility of attaining it through a cyclical set of pure-strategy actions, cannot be applied here since that result requires all players to be very patient.

Proof:

Define a punishment regime (à la Fudenberg and Maskin, 1986) where both players play their respective minimaxing payoffs (\hat{V}_M, \hat{V}_P) once a deviation occurs.

Play this for z number of stages, enough to fully remove whatever the deviant has gained, then both move back to the contract path. If there is any deviation while in the punishment regime, then restart the punishment regime.

We conclude from Lemmas 1-3 that any deviation could only be made most rewarding for M during the very first stage of cooperation, while for P , it is during the last stage of cooperation. This means that under a strategy profile $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$, the punishment regime could independently set in on the $(k+1)^{th}$ stage and on the $(k+j)^{th}$ stage for M and P , respectively. As before, to guarantee a no-deviation game scenario, each player's payoff over the entire game must be at least as much as their respective highest (entire-game) deviation payoffs. Hence,

$$\Pi_M^{\langle k,j \rangle} = \frac{(\bar{V} - V_c)(1 - \delta_M^k)}{1 - \delta_M^{k+j}} + V_c \geq \bar{V}(1 - \delta_M^{k+1}) + \varphi_M \delta_M^{k+1}, \quad (14)$$

$$\text{where } \varphi_M = (1 - \delta_M^z) \hat{V}_M + \delta_M^z \Pi_M^{\langle k,j \rangle}$$

$$\Pi_P^{\langle k,j \rangle} = \frac{(\underline{V} - V_c)(1 - \delta_P^k)}{1 - \delta_P^{k+j}} + V_c \geq \underline{V} + (V_c - \underline{V}) \delta_P^k + (\bar{V} - \bar{V} \delta_P - V_c) + \varphi_P \delta_P^{k+j} \quad (15)$$

$$\text{where } \varphi_P = (1 - \delta_P^z) \hat{V}_P + \delta_P^z \Pi_P^{\langle k,j \rangle}$$

In case a deviation occurs, φ_M and φ_P will be the respective discounted average payoffs of the two players during punishment regime, as computed based on the mechanics described in the first paragraph of this proof. Notice that whenever a player deviates during this regime, z increases since punishment regime starts again. And since both φ_M and φ_P decline continuously as z increases, both players will find no gain from deviating (or from not punishing) at this phase, thus, making the punishment regime binding.

By the minimum level of z that satisfies equations (14) and (15), we are assured that any gain from a one-shot deviation is neutralized and therefore not worth taking in the end. Now, think of a largest one-shot deviation that requires an almost infinite number of stages, z , to wipe out the gain that the deviant has obtained. This pushes down φ_M and φ_P to their respective limit value of \hat{V}_M and

\hat{V}_P (i.e. since $\lim_{z \rightarrow \infty} \varphi_i = \hat{V}_i$). By substituting \hat{V}_M and \hat{V}_P respectively to equations (14) and (15), we reached the same no-deviation conditions given in Proposition 2 and can show subsequently the similar conditions, as in Theorem 1(b), that generate individually rational outcomes. And since this result analogously applies to profile $(\sigma_M^{(j,k)}, \sigma_P^{(j,k)})$, we have completed the proof. ■

8. Final Remarks

The PTTS presented here are not the only types of TTS that are sustainable in perfect equilibrium. Other tractable forms of TTS, though maybe quite complex in structure, may still prove to generate sets of equilibrium payoff (an escalating contract path $(\bar{V}, V_c, \bar{V}, \bar{V}, V_c, V_c, \bar{V}, \bar{V}, \bar{V}, V_c, V_c, V_c, \dots)$ is one example). Moreover, even within the realm of cyclical TTS, certain structures that are different from the presented PTTS may also generate sets of equilibrium payoffs. The problem, however, is that some of them may not have a monotonic property (as in Lemma 3) which makes it difficult to characterize the timing of the highest-yielding possible deviation. Consider the strategy profile with a recurrent contract structure $(\bar{V}, \bar{V}, V_c, \bar{V}, V_c, V_c, \bar{V}, \bar{V}, V_c)$. For some values of δ_M and δ_P , it is possible for M to have its highest temptation on the 5th period (instead of 3rd) while for P on the 6th period (instead of 9th). Therefore, when one is presented with a long unsystematic contract path that is infinitely repeated, the greatest possible temptation to deviate may lie somewhere in the middle of the contract regime which would be laborious to characterize. In the end, our treatment of equilibrium outcomes for TTS in this paper is not exhaustive and is limited only to simple periodic strategies.

Furthermore, our study is confined only to two-player games. It would still be possible to find equilibrium payoffs in an n -player case, provided that a periodic contract that exhibits monotonic payoff streams is adopted (although not the only means). However, the characterization of perfect equilibria may prove to be elusive as it may require a more sophisticated punishment system when there is more than one impatient player in a game. In this scenario, it seeks to determine how the number of impatient players influence the equilibrium outcomes of an n -player

game, given the players' varying discount factors. We leave these questions at this moment open for further research.

References:

Abreu, D. (1988), "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica*, vol.56, no. 2, pp.383-396.

_____ (1986), "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory*, 39:191:225.

Aumann, R. and L. Shapley (1976), "Long Term Competition: A Game Theoretic Analysis", mimeo, Hebrew University.

Benoit, J.P. and V. Krishna (1985), "Finitely Repeated Games", *Econometrica*, vol.53, no.4, pp.905-922.

Chen, B. (2008), "On Effective Minimax Payoffs and Unequal Discounting", *Economic Letters*, 100: 105-107.

Friedman, J. (1971), "A Noncooperative Equilibrium for Supergames", *Review of Economic Studies*, 38:1-12.

Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Complete Information", *Econometrica*, vol. 54, no.3, pp.533-554.

_____ (1991), "On the Dispensability of Public Randomization in Discounted Repeated Games". *Journal of Economic Theory*, 53(2):428-438.

Fudenberg, D., D. Kreps, E. Maskin (1990), "Repeated Games with Long-run and Short-run Players", *Review of Economic Studies*, 57:557-573.

Gueron, Y., T. Lamadon, and C. Thomas, 2010, "On the Folk Theorem with One-Dimensional Payoffs and Different Discount Factors", *unpublished*.

Lehrer, E. and A. Pauzner (1999), "Repeated Games with Differential Time Preferences", *Econometrica*, vol. 67, no.2, pp.393-412.

Rubinstein, A (1979), "Equilibrium in Supergames with the Overtaking Criterion", *Journal of Economic Theory*, 21: 1-9.

Salonen, H. and H. Vartiainen (2008), "Valuating Payoff Streams under Unequal Discount Factors", *Economic Letters*, 99: 595-598.

Wen, Quan (1994), "The 'Folk Theorem' for Repeated Games with Complete Information". *Econometrica*, vol.62, pp. 949-954.

Appendix:

Proposition 2: (Equivalence) For any $k, j \in \mathbb{Z}^+$, the PTTS $(\sigma_M^{\langle k,j \rangle}, \sigma_P^{\langle k,j \rangle})$ and $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$ constitute the same range of values of δ_M and δ_P that can support the no-deviation condition during the contract regime of the repeated game. These values are defined by the following conditions:

$$\text{For player } M: \quad \delta_M^j \geq \frac{\tilde{\delta} - \delta_M^k}{\tilde{\delta} - \delta_M^{k+1}}$$

$$\text{For player } P: \quad \delta_P^j \leq \delta_P^{-k} + A \frac{1 - \delta_P^k}{\delta_P^{k-1}(\delta_P - \tilde{\delta})}, \text{ where } A < 0$$

Proof: We prove this directly by showing that the simplified form of $\Pi_i^{\langle k,j \rangle} \geq D_i^{\langle k,j \rangle}$ and $\Pi_i^{\langle j,k \rangle} \geq D_i^{\langle j,k \rangle}$ are the same for each i .

$$(A) \quad \Pi_i^{\langle k,j \rangle} \geq D_i^{\langle k,j \rangle}:$$

a1) For player M :

$$\begin{aligned} \Pi_M^{\langle k,j \rangle} &= \frac{(\bar{V} - V_c)(1 - \delta_M^k)}{1 - \delta_M^{k+j}} + V_c \geq \bar{V} - (\bar{V} - \hat{V})\delta_M^{k+1} = D_M^{\langle k,j \rangle} \\ \Rightarrow \frac{(\bar{V} - V_c)(1 - \delta_M^k)}{(\bar{V} - V_c) - (\bar{V} - \hat{V})\delta_M^{k+1}} &\geq 1 - \delta_M^{k+j} \quad (\text{The denominator is always positive since} \\ \frac{(\bar{V} - V_c)}{(\bar{V} - \hat{V})} &= \tilde{\delta} > \delta_M^{k+1}) \\ \Rightarrow \delta_M^{k+j} &\geq 1 - \frac{\tilde{\delta}(1 - \delta_M^k)}{(\tilde{\delta} - \delta_M^{k+1})} \Rightarrow \delta_M^j \geq \frac{\tilde{\delta} - \delta_M^k}{\tilde{\delta} - \delta_M^{k+1}}. \end{aligned}$$

a2) For player P :

$$\begin{aligned} \Pi_P^{\langle k,j \rangle} &= \frac{(V - V_c)(1 - \delta_P^k)}{1 - \delta_P^{k+j}} + V_c \geq V - (V - V_c)\delta_P^k + (\bar{V} - V_c)\delta_P^{k+j-1} - (\bar{V} - \hat{V})\delta_P^{k+j} = D_P^{\langle k,j \rangle} \\ \Rightarrow \frac{(V - V_c)(1 - \delta_P^k)}{1 - \delta_P^{k+j}} &\geq (V - V_c)(1 - \delta_P^k) + (\bar{V} - V_c)\delta_P^{k+j-1} - (\bar{V} - \hat{V})\delta_P^{k+j} \end{aligned}$$

$$\begin{aligned}
&\Rightarrow \frac{(V-V_c)(1-\delta_P^k)}{(\bar{V}-\hat{V})(1-\delta_P^{k+j})} \geq \frac{(V-V_c)}{(\bar{V}-\hat{V})}(1-\delta_P^k) + \tilde{\delta}\delta_P^{k+j-1} - \delta_P^{k+j} \\
&\Rightarrow \frac{(V-V_c)(1-\delta_P^k)}{(\bar{V}-\hat{V})(1-\delta_P^{k+j})} \geq \frac{\tilde{\delta}-\delta_P}{\delta_P} \Rightarrow \delta_P^{k+j} \leq 1 + \frac{(V-V_c)}{(\bar{V}-\hat{V})} \frac{\delta_P(1-\delta_P^k)}{(\delta_P-\tilde{\delta})} \\
&\Rightarrow \delta_P^j \leq \delta_P^{-k} + A \frac{1-\delta_P^k}{\delta_P^{k-1}(\delta_P-\tilde{\delta})}, \text{ where } A = \frac{V-V_c}{\bar{V}-\hat{V}} < 0.
\end{aligned}$$

(B) $\Pi_i^{\langle j,k \rangle} \geq D_i^{\langle j,k \rangle}$:

First, we note that in a no-deviation scenario, the strategy $(\sigma_M^{\langle j,k \rangle}, \sigma_P^{\langle j,k \rangle})$ yields:

$$\Pi_M^{\langle j,k \rangle} = \frac{(V_c-\bar{V})(1-\delta_M^j)}{(1-\delta_M^{j+k})} + \bar{V} \text{ and } \Pi_P^{\langle j,k \rangle} = \frac{(V_c-\underline{V})(1-\delta_P^j)}{(1-\delta_P^{j+k})} + \underline{V}.$$

Then, we write as follows the discounted payoff over the entire game of a one-shot deviation scenario. Note from the analogue of Lemma 2 that starts with cooperation that it is never profitable to deviate during the stages of tolerance, thus,

$$\begin{aligned}
D_M^{\langle j,k \rangle} &= (1-\delta_M) \left(\sum_{t=1}^q V_c \delta_M^{t-1} + \bar{V} \delta_M^q + \sum_{t=q+2}^{\infty} \hat{V} \delta_M^{t-1} \right) \text{ and} \\
D_P^{\langle j,k \rangle} &= (1-\delta_P) \left(\sum_{t=1}^r V_c \delta_P^{t-1} + \bar{V} \delta_P^r + \sum_{t=r+2}^{\infty} \hat{V} \delta_P^{t-1} \right),
\end{aligned}$$

where q and r are again the numbers of stages given to cooperation by M and P , respectively, just before defecting in the next stage. From Lemma 3, observe that $D_i^{\langle j,k \rangle} = \theta_i(\cdot)$ when k is set to 0, for both i . This shows that $D_i^{\langle j,k \rangle}$ is also monotonic in q and r , in a same manner specified in Lemma 3, thus:

(b1) For player M , the highest deviation payoff occurs when $q = 0$:

$$\begin{aligned}
\Pi_M^{\langle j,k \rangle} &= \frac{(V_c-\bar{V})(1-\delta_M^j)}{(1-\delta_M^{j+k})} + \bar{V} \geq \bar{V}(1-\delta_M) + \hat{V}\delta_M = D_M^{\langle j,k \rangle} \\
&\Rightarrow \tilde{\delta}(1-\delta_M^j) \leq \delta_M(1-\delta_M^{k+j}) \Rightarrow \delta_M^{k+j+1} \leq \delta_M - \tilde{\delta} + \tilde{\delta}\delta_M^j \\
&\Rightarrow \delta_M^j(\delta_M^{k+1} - \tilde{\delta}) \leq \delta_M - \tilde{\delta} \Rightarrow \delta_M^j \geq \frac{\tilde{\delta}-\delta_M}{\tilde{\delta}-\delta_M^{k+1}}.
\end{aligned}$$

(b2) For player P , the highest deviation payoff occurs when $r = j - 1$:

$$\begin{aligned}
\Pi_P^{\langle j,k \rangle} &= \frac{(V_c-\underline{V})(1-\delta_P^j)}{(1-\delta_P^{j+k})} + \underline{V} \geq V_c + (\bar{V}-V_c)\delta_P^{j-1} - (\bar{V}-\hat{V})\delta_P^j = D_P^{\langle j,k \rangle} \\
&\Rightarrow (V_c-\underline{V})(1-\delta_P^j) \geq \left[(V_c-\underline{V}) + (\bar{V}-V_c)\delta_P^{j-1} - (\bar{V}-\hat{V})\delta_P^j \right] (1-\delta_P^{j+k}) \\
&\Rightarrow \delta_P^j(V_c-\underline{V})(\delta_P^k-1) \geq \delta_P^{j-1}(\bar{V}-V_c-\bar{V}\delta_P+\hat{V}\delta_P)(1-\delta_P^{j+k}) \\
&\Rightarrow \delta_P \frac{(V-V_c)}{(\bar{V}-\hat{V})}(1-\delta_P^k) \geq (\delta_P-\tilde{\delta})(\delta_P^{j+k}-1) \Rightarrow \delta_P^{j+k} \leq 1 + A\delta_P \frac{(1-\delta_P^k)}{(\delta_P-\tilde{\delta})} \\
&\Rightarrow \delta_P^j \leq \delta_P^{-k} + A \frac{(1-\delta_P^k)}{\delta_P^{k-1}(\delta_P-\tilde{\delta})}, \text{ where } A < 0.
\end{aligned}$$

Comparing the results of (a1) with (b1) and (a2) with (b2), we conclude that the conditions $\Pi_i^{(j,k)} \geq D_i^{(j,k)}$ and $\Pi_i^{(k,j)} \geq D_i^{(k,j)}$ constitute the same range of values for δ_M and δ_P for every given k and j . \blacksquare

Lemma 4. *The real roots δ_M and δ_P of the equations $\delta_M^j = \frac{\tilde{\delta} - \delta_M}{\tilde{\delta} - \delta_M^{k+1}}$ and $\delta_P^j = \delta_P^{-k} + A \frac{1 - \delta_P^k}{\delta_P^{k-1}(\delta_P - \tilde{\delta})}$, respectively, that exist and belong to the interval $(0, 1)$, behave in the following manner with respect to k and j , for any $k, j \in \mathbb{Z}^+$:*

$$(i) \frac{\partial \delta_M}{\partial k} < 0 \text{ and } \frac{\partial \delta_M}{\partial j} > 0 \quad (ii) \frac{\partial \delta_P}{\partial k} > 0 \text{ and } \frac{\partial \delta_P}{\partial j} < 0$$

Proof:

(i) Let the first equation be redefined as an implicit function $F(k, j, \delta_M) := \frac{\delta_M - \delta_M^{k+j+1}}{1 - \delta_M^j} = \tilde{\delta}$. Then, $\frac{\partial \delta_M}{\partial k} = -\frac{\partial F / \partial k}{\partial F / \partial \delta_M}$ and $\frac{\partial \delta_M}{\partial j} = -\frac{\partial F / \partial j}{\partial F / \partial \delta_M}$, where $\frac{\partial F}{\partial \delta_M} \neq 0$ for any $\delta_M \in (0, 1)$ and $k, j \in \mathbb{Z}^+$.

By differentiating, we have

$$\begin{aligned} \frac{\partial F}{\partial \delta_M} &= \frac{(1 - \delta_M^j)(1 - (k + j + 1)\delta_M^{k+j}) + j(1 - \delta_M^{k+j})\delta_M^j}{(1 - \delta_M^j)^2} \\ &= \frac{\delta_M^{k+j} [(k + 1)(\delta_M^j - 1) - j] + (j - 1)\delta_M^j + 1}{(1 - \delta_M^j)^2} \end{aligned}$$

We start by setting $k = j = 1$ which gives us $\frac{\partial F}{\partial \delta_M} = \frac{2\delta_M^3 - 3\delta_M^2 + 1}{(1 - \delta_M)^2} > 0$, for any $\delta_M \in (0, 1)$. We show that the numerator, denoted as z , further increases away from zero when either k or j increases. First, $\frac{\partial z}{\partial k} = \delta_M^{k+j} [(\delta_M^j - 1) - ((k + 1)(1 - \delta_M^j) + j) \ln \delta_M]$. For this expression to be positive, it must be that $(\delta_M^j - 1) > ((k + 1)(1 - \delta_M^j) + j) \ln \delta_M$, or equivalently,

$$\frac{-1}{\ln \delta_M} < (k + 1) + \frac{j}{(1 - \delta_M^j)}$$

Note that $\ln \delta_M < 0$ for all $\delta_M \in (0, 1)$, and that the right hand side is least when $k = j = 1$, i.e. $\frac{-1}{\ln \delta_M} < 2 + \frac{1}{(1 - \delta_M)}$, which is always true for all $\delta_M \in (0, 1)$. Thus, increasing k only increases the right hand side, making $\frac{\partial z}{\partial k} > 0$. Next, we

show that $\frac{\partial z}{\partial j} > 0$. Observe that,

$$\begin{aligned}
z &= \delta_M^{k+j} [(k+1)(\delta_M^j - 1) - j] + (j-1)\delta_M^j + 1 \\
&> \delta_M^{k+j} [(k+1)(\delta_M^j - 1) - j] + (j-1)\delta_M^{k+j} + 1 \\
&> \delta_M^{k+j} [(k+1)(\delta_M^j - 1) - 1] + 1 \\
&> 0 \text{ when } k = j = 1, \text{ for all } \delta_M \in (0, 1).
\end{aligned}$$

Although the first term is always negative, it approaches zero as j increases. Thus, $z > 0$ for any $k, j \in \mathbb{Z}^+$ and $\delta_M \in (0, 1)$, which implies that $\frac{\partial F}{\partial \delta_M} > 0$.

Now, since $\ln \delta_M < 0$, for all $\delta_M \in (0, 1)$, we see that $\frac{\partial F}{\partial k} = \frac{-\delta_M^{k+j+1} \ln \delta_M}{1 - \delta_M^j} > 0$ and $\frac{\partial F}{\partial j} = \frac{\delta_M^{j+1}(1 - \delta_M^k) \ln \delta_M}{(1 - \delta_M^j)^2} < 0$, for any $k, j \in \mathbb{Z}^+$ and $\delta_M \in (0, 1)$. Following the formula above, we obtain $\frac{\partial \delta_M}{\partial k} < 0$ and $\frac{\partial \delta_M}{\partial j} > 0$.

(ii) Let the second equation be redefined as an implicit function $G(k, j, \delta_P) := \delta_P \left(1 + \frac{A(1 - \delta_P^k)}{(1 - \delta_P^{k+j})}\right) = \tilde{\delta}$. Then, $\frac{\partial G}{\partial \delta_P} = 1 + \frac{A\hat{z}}{(1 - \delta_P^{k+j})^2}$, where $\hat{z} = (1 - \delta_P^{k+j})(1 - (k+1)\delta_P^k) + (k+j)(1 - \delta_P^k)\delta_P^{k+j}$. We will show that $\frac{\partial G}{\partial \delta_P} > 0$ for any $k, j \in \mathbb{Z}^+$. First, observe that

$$\begin{aligned}
\hat{z} &= 1 - (k+1)\delta_P^k + (k+j-1)\delta_P^{k+j} - (j-1)\delta_P^{2k+j} \\
&> 1 + k\delta_P^{k+j} - (k+1)\delta_P^k = 1 + \delta_P^k(k\delta_P^j - (k+1)) \\
&> 0 \text{ for } k = j = 1.
\end{aligned}$$

As k increases, \hat{z} remains positive and approaches 1; while as j increases, $\hat{z} > 0$ for as long as $\delta_P < \left(\frac{1}{k+1}\right)^{\frac{1}{k}}$, otherwise if $\hat{z} \leq 0$, we are done with $\frac{\partial G}{\partial \delta_P} > 0$ since A is negative. Thus, when $\hat{z} > 0$, $\frac{\partial G}{\partial \delta_P} > 0$ iff $A \geq -\frac{(1 - \delta_P^{k+j})^2}{\hat{z}}$. To determine the least lower bound of A , see that (a) as $k \rightarrow \infty$ for any j , $A \geq -1$; (b) as $j \rightarrow \infty$, the bound is least when $k = 1$, *i.e.* $A \geq -\frac{1}{1 - 2\delta_P}$; and (c) for $k = j = 1$, $A \geq -(1 + \delta_P)^2$. By (b), $A \rightarrow -\infty$ as $\delta_P \rightarrow (1/2)_-$ and so $\frac{\partial G}{\partial \delta_P} > 0$ for any $k, j \in \mathbb{Z}^+$ and for $A < 0$. Next, we see that $\frac{\partial G}{\partial k} = \frac{A\delta_P^{k+1}(\delta_P^j - 1) \ln \delta_P}{(1 - \delta_P^{k+j})^2} < 0$ and $\frac{\partial G}{\partial j} = \frac{A\delta_P^{k+j+1}(1 - \delta_P^k) \ln \delta_P}{(1 - \delta_P^{k+j})^2} > 0$ for all $\delta_P \in (0, 1)$ and $k, j \in \mathbb{Z}^+$. Hence, from the analogous formula in (i), we have $\frac{\partial \delta_P}{\partial k} > 0$ and $\frac{\partial \delta_P}{\partial j} < 0$. \blacksquare

Lemma 5: For any given $k, j \in \mathbb{Z}^+$, $\alpha \in \mathbb{Z}^+ \setminus \{1\}$, $\delta_M \in (\underline{\delta}_M, \tilde{\delta})$, and $\delta_P \in$

$(\delta_P, 1) :$

- (i) $\Pi_M^{\langle k,j \rangle} > \Pi_M^{\langle j,k \rangle}$
- (ii) $\Pi_P^{\langle j,k \rangle} > \Pi_P^{\langle k,j \rangle}$
- (iii) $\Pi_M^{\langle j,k \rangle} > \Pi_M^{\langle \alpha j, \alpha k \rangle}$ and $\Pi_M^{\langle k,j \rangle} < \Pi_M^{\langle \alpha k, \alpha j \rangle}$
- (iv) $\Pi_P^{\langle j,k \rangle} > \Pi_P^{\langle j, \alpha k \rangle}$ and $\Pi_P^{\langle k,j \rangle} > \Pi_P^{\langle \alpha k, j \rangle}$

Proof:

- (i) Suppose $\Pi_M^{\langle k,j \rangle} - \Pi_M^{\langle j,k \rangle} \leq 0$. Then, $\frac{(\bar{V}-V_c)(1-\delta_M^k)}{1-\delta_M^{k+j}} + V_c - \frac{(V_c-\bar{V})(1-\delta_M^j)}{(1-\delta_M^{j+k})} - \bar{V} \leq 0$.
 $\Rightarrow (\bar{V} - V_c) \left(\frac{(1-\delta_M^k)}{(1-\delta_M^{k+j})} + \frac{(1-\delta_M^j)}{(1-\delta_M^{j+k})} - 1 \right) \leq 0 \Rightarrow (\bar{V} - V_c)(1-\delta_M^j)(1-\delta_M^k) \leq 0$.

Since $\bar{V} - V_c > 0$ and that for any $\delta_M^k, \delta_M^j \in (0, 1)$ for any finite $k, j \in \mathbb{Z}^+$, the above inequality is a contradiction, thus $\Pi_M^{\langle k,j \rangle} - \Pi_M^{\langle j,k \rangle} > 0$.

- (ii) Suppose $\Pi_P^{\langle j,k \rangle} - \Pi_P^{\langle k,j \rangle} \leq 0$. Then, $\frac{(V_c-\underline{V})(1-\delta_P^j)}{(1-\delta_P^{j+k})} + \underline{V} - \frac{(V_c-\underline{V})(1-\delta_P^k)}{(1-\delta_P^{k+j})} - V_c \leq 0$.
 $\Rightarrow (V_c - \underline{V}) \left(\frac{(1-\delta_P^j)}{(1-\delta_P^{j+k})} + \frac{(1-\delta_P^k)}{(1-\delta_P^{k+j})} - 1 \right) \leq 0 \Rightarrow (V_c - \underline{V})(1-\delta_P^j)(1-\delta_P^k) \leq 0$.

Since $V_c > \underline{V}$ and with the similar argument as (i) above, we have a contradiction. Therefore, $\Pi_P^{\langle j,k \rangle} - \Pi_P^{\langle k,j \rangle} > 0$.

- (iii) Suppose $\Pi_M^{\langle j,k \rangle} \leq \Pi_M^{\langle \alpha j, \alpha k \rangle}$. Then, we have:

$$\begin{aligned} \Pi_M^{\langle j,k \rangle} &= \frac{(V_c-\bar{V})(1-\delta_M^j)}{(1-\delta_M^{j+k})} + \bar{V} \leq \frac{(V_c-\bar{V})(1-\delta_M^{\alpha j})}{(1-\delta_M^{\alpha(j+k)})} + \bar{V} = \Pi_M^{\langle \alpha j, \alpha k \rangle} \\ &\Rightarrow \frac{(1-\delta_M^j)}{(1-\delta_M^{j+k})} \geq \frac{(1-\delta_M^{\alpha j})}{(1-\delta_M^{\alpha(j+k)})} \text{ since } V_c - \bar{V} < 0 \\ &\Rightarrow \frac{(1-\delta_M^j)(1+\delta_M^j+\delta_M^{2j}+\dots+\delta_M^{(\alpha-1)j})}{(1-\delta_M^{j+k})} \geq \frac{(1-\delta_M^{\alpha j})(1+\delta_M^{\alpha j}+\delta_M^{2\alpha j}+\dots+\delta_M^{(\alpha-1)\alpha j})}{(1-\delta_M^{\alpha(j+k)})} \\ &\Rightarrow \delta_M^j(1-\delta_M^k) + \delta_M^{2j}(1-\delta_M^{2k}) + \dots + \delta_M^{(\alpha-1)j}(1-\delta_M^{(\alpha-1)k}) \leq 0, \text{ which is} \\ &\text{a contradiction since all terms are positive for } \delta_M \in (0, 1), k, j \in \mathbb{Z}^+, \text{ and } \alpha \in \mathbb{Z}^+ \setminus \{1\}. \text{ Hence, } \Pi_M^{\langle j,k \rangle} > \Pi_M^{\langle \alpha j, \alpha k \rangle}. \end{aligned}$$

Similarly, suppose $\Pi_M^{\langle k,j \rangle} \geq \Pi_M^{\langle \alpha k, \alpha j \rangle}$. Then,

$$\begin{aligned} \Pi_M^{\langle k,j \rangle} &= \frac{(\bar{V}-V_c)(1-\delta_M^k)}{1-\delta_M^{k+j}} + V_c \geq \frac{(\bar{V}-V_c)(1-\delta_M^{\alpha k})}{1-\delta_M^{\alpha(k+j)}} + V_c = \Pi_M^{\langle \alpha k, \alpha j \rangle} \\ &\Rightarrow \frac{(1-\delta_M^k)(1+\delta_M^k+\delta_M^{2k}+\dots+\delta_M^{(\alpha-1)k})}{(1-\delta_M^{k+j})} \geq \frac{(1-\delta_M^{\alpha k})(1+\delta_M^{\alpha k}+\delta_M^{2\alpha k}+\dots+\delta_M^{(\alpha-1)\alpha k})}{(1-\delta_M^{\alpha(k+j)})} \\ &\Rightarrow \delta_M^k(1-\delta_M^j) + \delta_M^{2k}(1-\delta_M^{2j}) + \dots + \delta_M^{(\alpha-1)k}(1-\delta_M^{(\alpha-1)j}) \leq 0, \text{ also a contra-} \\ &\text{diction.} \end{aligned}$$

- (iv) Suppose $\Pi_P^{\langle j, \alpha k \rangle} \geq \Pi_P^{\langle j,k \rangle}$. Then $\Pi_P^{\langle j, \alpha k \rangle} = \frac{(V_c-\underline{V})(1-\delta_P^j)}{(1-\delta_P^{j+\alpha k})} + \underline{V} \geq \frac{(V_c-\underline{V})(1-\delta_P^j)}{(1-\delta_P^{j+k})} + \underline{V} = \Pi_P^{\langle j,k \rangle}$

$\Rightarrow \frac{1}{(1-\delta_P^{j+\alpha k})} \geq \frac{1}{(1-\delta_P^{j+k})} \Rightarrow \delta_P^{j+\alpha k} \geq \delta_P^{j+k}$, a contradiction for all $\delta_P \in (0, 1)$,
 $k, j \in \mathbb{Z}^+$, and $\alpha \in \mathbb{Z}^+ \setminus \{1\}$. Hence, $\Pi_P^{\langle j, \alpha k \rangle} < \Pi_P^{\langle j, k \rangle}$.
 Suppose $\Pi_P^{\langle \alpha k, j \rangle} \geq \Pi_P^{\langle k, j \rangle}$. Then, $\Pi_P^{\langle \alpha k, j \rangle} = \frac{(V-V_c)(1-\delta_P^{\alpha k})}{(1-\delta_P^{\alpha k+j})} + V_c \geq \frac{(V-V_c)(1-\delta_P^k)}{(1-\delta_P^{k+j})} + V_c$
 $\Rightarrow (1-\delta_P^{\alpha k})(1-\delta_P^{k+j}) \leq (1-\delta_P^{\alpha k+j})(1-\delta_P^k)$, since $V-V_c < 0$
 $\Rightarrow (1-\delta_P^j)(\delta_P^k - \delta_P^{\alpha k}) \leq 0$, which is not true for any $\delta_P \in (0, 1)$ and
 $\alpha \in \mathbb{Z}^+ \setminus \{1\}$. ■