Repeated Games with Endogenous Discounting*

Asen Kochov   Yangwei Song†

March 12,2014

This paper studies infinitely repeated games in which discount factors can depend on actions. One of the main results is that in any efficient equilbrium of a repeated prisoners' dilemma game, the players must eventually cooperate. Depending on the parameters of the model, cooperation can be either intratemporal or intertemporal. The result suggests that the multiplicity of efficient equilibria, traditionally associated with repeated games, is an artefact of the time-additive preference specification in which the rate of discount is constant.

KEYWORDS: Repeated games, efficiency, folk theorem, endogenous discount factors.

---

*We are grateful to Paulo Barelli, Hari Govindan, and Takuo Sugaya for their helpful comments.

†Department of Economics, University of Rochester, Rochester, NY 14627. E-mail: asen.kochov@rochester.edu, ysong19@z.rochester.edu

1 Introduction

Strong restrictions on the structure of intertemporal preferences are a common feature in the study of repeated games. In fact, most of the literature assumes that preferences can be represented by an additive payoff function with a constant rate of time preference. This specification has limited descriptive or normative appeal. Its primary advantage is analytic tractability. This paper considers a more general class of intertemporal preferences introduced by Uzawa [14]. Specifically, the discounted sum of payoffs is defined recursively as

$$v_i(a^0, a^1, ...) = g_i(a^0) + \beta_i(a^0) v_i(a^1, a^2, ...) \tag{1}$$

where $g_i(a)$ is player i's stage payoff from an action profile a and $\beta_i(a)$ is the player's discount factor as a function of that action. Repeated games in which intertemporal preferences take this form are referred to as games with endogenous discounting or ED games for short.

The paper achieves two objectives. The first is to establish a folk theorem for ED games. In this regard, the familiar conclusion that every sequentially rational outcome can arise in an equilibrium of a repeated game is confirmed. Below we describe some of the main challenges encountered in proving a folk theorem. The second objective is to investigate how efficiency restricts the set of equilibrium outcomes. While some of the findings apply more generally, in this part of the paper attention is mainly limited to a repeated prisoners' dilemma game. The potentially surprising conclusion is that, for a natural subclass of the preferences we consider, any efficient equilibrium leads to an eventually unique outcome. The result suggests that the multiplicity of efficient equilibria, traditionally associated with repeated interactions, is an artefact of the time-additive model. The rest of the introduction describes how we specialize the preferences in (2) and the precise meaning of uniqueness.

In the literature on endogenous discounting, it is common to assume that discount factors are a strictly monotonic function of the underlying outcomes. This is also the assumption under which we investigate whether equilibrium outcomes are unique. There are two cases to consider. One is that the marginal impatience of each player i, $1 - \beta_i(a)$, decreases the more desirable he finds the constant path $(a, a, ...)$. Increasing marginal impatience is the obvious polar case. Interestingly, the merits of each case have been debated going back to the classical works of Fisher et al. [7, p.72] and Friedman [8, p.30]. Epstein [5, 6] provides a comprehensive summary of the arguments that have been made. In this paper, we do not take sides in this debate. The two cases lead to different uniqueness results. Since the driving forces behind them are also quite different, we investigate each case in turn.

To illustrate the uniqueness results, consider a repeated prisoners' dilemma game. As usual, let C stand for 'cooperate' and D for 'defect'. An interesting consequence of the more flexible class of preferences we study is the need to distinguish between two forms of dynamic cooperation. In particular, say that cooperation is *intratemporal* if (C, C) is played in every period. Cooperation is *intertemporal* if the players alternate between (D, C) and (C, D), i.e., between their most preferred outcomes. With these definitions in place, consider the case when marginal impatience is increasing. The paper shows that, after some period, the play path in any efficient equilibrium is one of dynamic cooperation. Furthermore, whether cooperation is intratemporal or intertemporal is independent of the equilibrium being played. The outcome is fully determined by the specification of preferences. Consequently, all efficient equilibria of the repeated game result, eventually, in a unique play path. If marginal impatience is decreasing, the uniqueness result is stronger. In any efficient equilibrium, cooperation is intratemporal and starts immediately.

It is important to emphasize that, in the standard time-additive model, intertemporal cooperation is never efficient. Therefore, the more general class of preferences we consider do not simply restrict the set of outcomes that can arise in an efficient equilibrium. If marginal impatience is increasing, they can also generate different, potentially interesting dynamics.

In formulating a folk theorem for ED games, one of the main problems posed by adopting the more general utility specification in (1) is that a change in the rate of time preference may change the minmax strategies against each player. In general, one cannot therefore keep a sequentially rational outcome fixed and at the same time let discount factors converge to one. Our solution to this problem involves several steps. A preliminary result shows that the search for minmax strategies can be restricted to strategies that are history independent. The result extends the well-known fact that in standard repeated games it is sufficient to look at minmax strategies in the stage game. The next step is to identify a path along which discount factors converge to one, yet the ranking of history-independent strategies, and therefore the minmax strategies against each player, remain invariant. The appeal of the proposed convergence path is not limited to the role it plays in proving a folk theorem. It is chosen so that the relative impatience, $\frac{1-\beta_i(a)}{1-\beta_j(a')}$, between any two action profiles and any two players, remains constant as discount factors converge to one.

A final difficulty arises in games with more than two players. As is well-known from Fudenberg and Maskin [9], a rich set of actions is then necessary to insure that players can be rewarded for carrying out the punishments against a player who deviates. As before, we need the rewarding strategies to be history independent so that they remain invariant as discount factors converge to one. Because the preferences in (2) are non-additive, however, we have not been able to find such strategies under the traditional full-dimensionality condition introduced in Fudenberg and Maskin [9]. A stronger condition is proposed instead.[1]

An axiomatic foundation for the preferences we consider is provided by Epstein [5]. He shows that a utility representation as in (1) exists if and only if behavior is stationary and random play paths are evaluated according to the expected-utility criterion. Two aspects of his result are worth emphasizing here. First, the representation in (1) is obtained precisely by relaxing time additivity. The latter is arguably the most problematic feature of the standard model. The result also shows that, unless one is willing to abandon the more appealing properties of stationarity and expected utility, there is no room to pursue generalizations.

One limitation should be acknowledged from the start. As discussed in Fudenberg and Tirole [11, p.21], there are games in which an efficient equilibrium may not be the most reasonable prediction of how a game is played. In the uniqueness results we obtain, this paper uses the 'Pareto refinement' without providing any formal justification. While the efficient equilibria we find for the prisoners' dilemma seem intuitive, more research is certainly needed to understand when imposing efficiency is appropriate from a positive standpoint.

2 Related Literature

Recursive preferences with endogenous discounting have been previously used in the literature on optimal growth. See Epstein [6], Lucas and Stokey [13], and the survey in

[1] We conjecture that this condition can be substantially relaxed. Since the problem is only tangential to the analysis of uniqueness, which is our main concern, we leave this to future work.

Backus et al. [1]. In departing from the standard model, one of the primary motivation of these papers is to escape the well-known 'immiseration' result of Becker [3]. The result states that only the most patient player can own capital in a steady state of the economy. This has long been viewed as an unappealing consequence of the time-additive specification. Both Epstein [6] and Lucas and Stokey [13] emphasize the role of increasing marginal impatience in insuring that there is a unique steady state and that the distribution of wealth is nondegenerate. The uniqueness results appear conceptually related to ours. However, we have not been able to find any formal connection. The continuous choice framework in those papers is different from the discrete games which are the focus of this paper. In addition, both Epstein [6] and Lucas and Stokey [13] impose additional, auxiliary assumptions on preferences which are not needed here. Finally, it should be mentioned that these papers do not investigate the case of decreasing marginal impatience.

Lehrer and Pauzner [12] provide a game-theoretic analogue of the immiseration result of Becker [3]. In their paper, discount factors are constant but heterogeneous across players. Efficiency then requires that the utility of the most patient player is eventually maximized. Unlike Becker [3], however, such an outcome need not be sequentially rational for the impatient player. This implies that a strategic setting such as the repeated prisoners' dilemma game has no efficient equilibrium. In this paper, heterogeneity in the players' rates of time preference can arise *endogenously* even when the players are a priori identical. This occurs along any path along which the players attain different outcomes. It will be clear from the discussion in Section 7.3 that such endogenous heterogeneity plays an important role in the analysis of this paper. The switch from exogenous to endogenous heterogeneity however produces substantially different results from those of Lehrer and Pauzner [12].

3 The Model

Time is discrete and varies over an infinite horizon $t \in \{0, 1, ...\}$. There is finite set of players $I = \{1, 2, ..., n\}$. In each period t, player i can choose a pure action in a finite set A_i. Mixed actions are denoted by $\alpha_i \in \Delta(A_i)$. To simplify the analysis, we permit public randomization: in each stage the players can condition their actions on an exogenous random variable. As is typical, we do not make the assumption explicit. A complete history up to some period t consists of all the past mixed actions, realized outcomes and public signals. We assume perfect monitoring: each player can condition his action at time t on the entire history. Let Σ_i denote the corresponding set of behavioral strategies for player $i \in I$ and let $\Sigma := \times_{i \in I} \Sigma_i$. A generic strategy profile is denoted as $\sigma = (\sigma_i)_i \in \Sigma$. A play path $\mathbf{a} = (a_0, a_1, ...,) \in A^\infty$ is a sequence of action profiles. Given a path $\mathbf{a} = (a_0, a_1, ...,) \in A^\infty$ and a time period $t \in T$, ${}_t\mathbf{a}$ denotes the continuation path $(a_t, a_{t+1}, ...)$ starting from period t. To describe player i's preferences, first define a utility function v_i on A^∞ as follows

$$v_i(\mathbf{a}) = g_i(a^0) + \beta_i(a^0)g_i(a^1) + \beta_i(a^0)\beta_i(a^1)g_i(a^2) + ... = g_i(a^0) + \beta_i(a^0)v_i({}_1\mathbf{a}) \qquad (2)$$

where $g_i : A \to \mathbb{R}$ is player i's stage payoff and $\beta_i : A \to (0,1)$ is his discount factor. Given (2), preferences can be extended to random strategy profiles in the usual manner. In particular, each strategy profile $\sigma \in \Sigma$ induces a probability distribution on A^∞. Abusing notation, we denote the induced measure by σ as well. Player i's expected payoff from a strategy profile σ is then $v_i(\sigma) := \mathbb{E}_\sigma v_i(\mathbf{a})$. Note that if each $\beta_i : A \to (0,1)$ is a constant function, one obtains the standard time-additive model with a constant rate of time preference.

By an equilibrium of an ED game, we always mean a subgame perfect equilibrium that induces a deterministic play path. An equilibrium is efficient if it yields payoffs on the Pareto frontier of the feasible set.

4 Equivalent ED Games

An important difference between ED games and standard repeated games has to be emphasized. For this, it is necessary to first understand the uniqueness of the utility representation in (2). For the moment, suppress the index i. Note from (2) that a preference relation on $\Delta(A^\infty)$ is uniquely determined by the tuple (β, g) where $\beta : A \to (0,1)$ and $g : A \to \mathbb{R}$. Epstein [5] shows that $(\beta, g), (\beta', g')$ represent the same preference relation on $\Delta(A^\infty)$ if and only if $\beta' = \beta$ and $g' = \alpha g + \gamma(1 - \beta)$ for some constants α, γ with $\alpha > 0$.[2] The key implication is that, unless $\gamma = 0$ or β is a constant function, the functions g and g' need not be cardinal or even monotone transformations of one another. Therefore, equivalent transformations of an ED game may not in general lead to equivalent stage

[2]Technically, the uniqueness of the function β requires that A is a topologically connected space. In this paper, we abstract away from the problem of uniqueness and simply assume that the β_i's are known.

games.

5 Minmax Strategies

One of the main problems in analyzing repeated games with endogenous discounting is that the minmax strategies of each player may change when discount factors change. In particular, the traditional result that allows us to identify minmax strategies in the repeated game with minmax strategies in the stage game is no longer available. The following lemma provides an appropriate generalization. Say that a strategy σ_i is constant or history independent if $\alpha_i \in \Delta_i$ is played in every history. Denote each such strategy by α_i^{con}.

Lemma 5.1. *For every player $i \in I$,*

$$\min_{\sigma_{-i} \in \times_{k \neq i} \Sigma_k} \max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \sigma_{-i}) = \min_{\alpha_{-i} \in \times_{k \neq i} \Delta A_k} \max_{\alpha_i \in \Delta A_i} v_i(\alpha_i^{con}, \alpha_{-i}^{con}).$$

Let $M^i \in \Sigma$ be the strategy profile such that $M^i_{-i} \in \times_{k \neq i} \Sigma_k$ are the minmax strategies against player i and $M^i_i \in \Sigma_i$ is his best reply. Normalize the utility functions so that $g_i(M^i) = 0$, for each $i \in I$. Given Lemma 5.1, assume that each strategy profile M^i consists of constant strategies.

For the rest of the analysis, it is useful to compute the payoffs from a constant strategy profile $\alpha^{con}, \alpha \in \Delta(A)$. When no confusion arises, we may write $v_i(\alpha)$ instead $v_i(\alpha^{con})$. For every $i \in I, \alpha \in \Delta(A)$, let $g_i(\alpha) := \sum_{a \in A} g_i(a)\alpha(a)$, and $\beta_i(\alpha) := \sum_{a \in A} \beta_i(a)\alpha(a)$. Note that each constant strategy induces an IID probability measure on A^∞. Consequently, the ex ante expected payoff from a constant strategy is equal to its expected payoff after any given history. Thus,

$$v_i(\alpha^{con}) = \mathbb{E}_\alpha[g_i(a) + \beta_i(a) v_i(\alpha^{con})] = g_i(\alpha) + \beta_i(\alpha) v_i(\alpha^{con}) \quad \Leftrightarrow \tag{3}$$

$$v_i(\alpha^{con}) = \frac{g_i(\alpha)}{1 - \beta_i(\alpha)}. \tag{4}$$

One can see from (8.3) that if discount factors are constant, as they are in the standard model, the ranking of constant strategies is determined entirely by the stage payoffs g_i. This is not true in general. In particular, a change in the rates of time preference may change the minmax strategies against each player and their respective security levels.

This complicates both the statement and the proof of our folk theorems.[3] As we now show however there are reasonable convergence paths along which the problem does not arise. Specifically, write the discount factor of each player i as follows

$$\beta_i(a) = 1 - \lambda(1 - \beta_i^0(a)) \tag{5}$$

where $0 < \lambda \leq 1$. In the rest of the paper, we will be concerned with equilibrium behavior as λ goes to zero. As mentioned in the introduction, note that the relative impatience, $\frac{1-\beta_i(a)}{1-\beta_j(a')}$, between any two action profiles and any two players, remains constant as discount factors approach 1.

Given the specification in (5), it is convenient to normalize the utility functions as follows

$$v_i(\mathbf{a}, \lambda) = \lambda \sum_{t=0}^{\infty} \prod_{\tau=0}^{t-1} \beta_i(a^\tau) g_i(a^t) = \lambda g_i(a^0) + \beta_i(a^0) v_i({}_1\mathbf{a}, \lambda) \tag{6}$$

where $\prod_{\tau=0}^{-1} \beta_i(a^\tau) = 1$. To emphasize that players' preferences may change as λ converges to zero, we sometimes make the dependence on λ explicit as in (6) above. However, and this is key for our analysis, the ranking of constant strategies is independent of λ for every player $i \in I$. Specifically, the expression in (8.3) becomes

$$v_i(\alpha^{con}, \lambda) = \frac{g_i(\alpha)}{1 - \beta_i^0(\alpha)} =: v_i(\alpha), \quad \forall \lambda \in (0, 1], \forall i \in I, \forall \alpha \in \Delta(A).$$

Using the normalized utility function, we can now define the following sets which will be used in the rest of the paper

$$\overline{V}(\lambda) := \{v(\sigma, \lambda) : \sigma \in \Sigma\} \qquad \text{and} \qquad V := \text{co}\{v(a) : a \in A\}.$$

where 'co' denotes the convex hull of a set. Thus, $\overline{V}(\lambda)$ is the set of feasible payoffs in the repeated game and V is the convex hull of all payoffs achievable by constant pure strategies. Also, define the set of feasible and strictly individually-rational payoffs:

$$V^*(\lambda) = \{(v_1, ..., v_n) \in \overline{V}(\lambda) : v_i > 0, i \in I\} \quad \text{and} \quad V^* = \{(v_1, ..., v_n) \in V : v_i > 0, i \in I\}.$$

[3] The statement of the folk theorem is complicated because the set of sequentially rational play paths changes with the discount factor. In stochastic games, where a similar problem arises, see e.g. Dutta [4], it is common to define sequential rationality using the asymptotic minmax levels. In this case, however, one cannot guarantee that efficient outcomes are achieved. The latter is important for our 'anti-folk' results in Section 7.

6 Folk Theorems

With endogenous discounting, the feasible set $\overline{V}(\lambda)$ may change as λ approaches 0 and the β_i's converge to one. For this reason, the folk theorem cannot be stated in terms of payoffs. Instead, we ask what paths $\mathbf{a} \in A^\infty$ can be sustained in a subgame perfect equilibrium of the repeated game. The common sufficient condition to establish Folk Theorem is the full dimensionality assumption, which was introduced by Fudenberg and Maskin [9]. Full dimensionality enables players to be rewarded for having carried out the minmax punishments. It provides incentives for the other players to punish the deviator in the rewarding phase of the punishment. With time-additive preferences, there exists a constant strategy that achieves any given payoff vector. Hence, full dimensionality ensures the existence of both rewarding payoffs and rewarding strategies. However, with endogenous discounting preferences, Lemma 6.2 implies that given some payoff profile, there may not exist a constant strategy to achieve it. There are two reasons why we require rewarding strategies to be constant. First, if the strategies are not constant, to obtain the same payoffs, the strategies may have to change as discount factors change. Second, using nonconstant strategies makes it difficult to verify individual rationality in each time period. Thus, we need a stronger assumption to ensure the existence of constant rewarding strategies. The following assumption requires that the set of action profiles should be rich enough.

Richness: For any $i \in I$, there exists action profiles a^i and $\tilde{a}^i$ such that $v_i(a^i) \leq 0 < v_i(\tilde{a}^i)$ and $v_j(a^i) > 0, v_j(\tilde{a}^i) \geq 0$.

Richness implies the usual full dimensionality condition. It is sufficient but not necessary for the Folk Theorem. It excludes the class of games in which all the actions but the minmax actions yield strictly positive payoffs for all the players. The following lemma states that under Richness, constant rewarding strategies exist.

Lemma 6.1. *Assume Richness. There exists $\overline{\epsilon}$ such that for any $0 < \epsilon < \overline{\epsilon}$, for any $i \in I$, we can find $\alpha^i \in \Delta A$ such that $v_i(\alpha^i) = \epsilon$ and $v_j(\alpha^i) > \epsilon$ for all $j \in I \setminus \{i\}$.*

Theorem 6.1. *Assume Richness. For any $\varepsilon > 0$ and for any path $\mathbf{a} \in A^\infty$ such that $v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$ for all $i \in I, t \in T, \lambda \in (0,1]$, there exists $\overline{\lambda} \in (0,1]$ such that for all $0 < \lambda < \overline{\lambda}$, $\mathbf{a}$ can be supported in a subgame perfect equilibrium.*

If we assume that $\beta_i(a) = \beta_j(a)$ for all $a \in A$ and $i, j \in I$, we can state a folk theorem in terms of payoffs as is typical in the literature. This is because the feasible set no longer depends on λ. Furthermore, any feasible payoff can be achieved by a constant mixed strategy. These results are summarized in the next lemma.

Lemma 6.2. *Suppose that $\beta_i(a) = \beta_j(a)$ for all $i, j \in I, a \in A$. Then, $V = \overline{V}(\lambda)$. Moreover, for any $v \in V$, there exists $\alpha \in \Delta A$ such that $v = v(\alpha)$.*

We emphasize that the assumption in Lemma 6.2 is extremely restrictive. In most games, the same action profiles $a \in A$ may give different players different outcomes. Consequently, the assumption rules out the possibility that the rate of time preference depends on outcomes. The assumption is consistent with models of endogenous discounting as developed by Becker and Mulligan [2] in which individuals make efforts to increase their patience.

Theorem 6.2. *Suppose that $\beta_i(a) = \beta_j(a)$ for all $i, j \in I, a \in A$. Suppose V^* has full dimension. Then, for any $v \in V^*$, there exists $\overline{\lambda} \in (0, 1]$ such that for all $0 < \lambda < \overline{\lambda}$ there exists a subgame perfect equilibrium of the infinitely repeated game in which player i receives payoff v_i.*

In two-player games, the richness condition in Theorem 6.1 and the full dimensionality condition in Theorem 6.2 are not needed. The proofs are similar to that of Theorem 1 in Fudenberg and Maskin [9]. It should also be mentioned that Fudenberg and Maskin [9, 10] extend their folk theorem to the case in which mixed actions are not observable and public randomization is not available. It is an open problem if such generalizations are valid for ED games.

7 Efficiency

In this section, we restrict attention to symmetric, two-player repeated games. The notion of symmetry needs clarification. As discussed in the introduction, we now assume that, for each player i, the discount factor $\beta_i(a)$ depends on the action profile only through player i's stage payoff. Symmetry then means that there is no *a priori* heterogeneity in this dependence. As a result, we are limiting attention to settings in which intertemporal trade and any heterogeneity in the rate of time preference can only arise endogenously, in the course of the game. The objective is to investigate if and when efficiency requires intertemporal trade and whether such allocations can be supported as equilibria of the game.

It is useful to recall some elementary facts about the Pareto frontier of the set of feasible payoffs $\overline{V}(\lambda)$. Since the set is convex, every point on the outer boundary maximizes a weighted sum of the players' payoffs. To characterize the frontier, it is furthermore enough to focus on the set of deterministic paths. In particular, for any pair of positive weights $\eta := (\eta_1, \eta_2) \in \mathbb{R}^2_+$, let $P(\lambda, \eta)$ be the set of efficient paths that solve the maximization problem

$$\max_{\mathbf{a} \in A^\infty} \eta_1 \sum_{t=0}^{\infty} \prod_{\tau=0}^{t-1} \beta_1(a^\tau) g_1(a^t) + \eta_2 \sum_{t=0}^{\infty} \prod_{\tau=0}^{t-1} \beta_2(a^\tau) g_2(a^t). \tag{7}$$

We now formalize two polar assumptions about how discount factors depend on action profiles.

Increasing Marginal Impatience (IMI): For each player i, for any $a, a' \in A$, $\frac{g_i(a)}{1-\beta_i(a)} > \frac{g_i(a')}{1-\beta_i(a')}$ if and only if $\beta_i(a) < \beta_i(a')$.

Decreasing Marginal Impatience (DMI): For each player i, for any $a, a' \in A$, $\frac{g_i(a)}{1-\beta_i(a)} > \frac{g_i(a')}{1-\beta_i(a')}$ if and only if $\beta_i(a) > \beta_i(a')$.

7.1 Increasing Marginal Impatience

Assume that marginal impatience is increasing. The focus of this section is a repeated prisoners' dilemma game. Let the action space A and the stage payoffs $(g_i)_{i \in I}$ be as in the matrix below,

	C	D
C	d, d	c, b
D	b, c	$0, 0$

Figure 1: The prisoners' dilemma

where, as usual, C stands for cooperate and D for defect. Because discount factors depend on outcomes only and this dependence is identical across players, we write $\beta^0(b) := \beta^0_1(D, C) = \beta^0_2(C, D)$ and similarly for all other outcomes in the matrix above. Together, the matrix in Figure 7.1 and the discount factors determine the utility functions $v_i : A^\infty \to \mathbb{R}$. Assume that $\frac{b}{1-\beta^0(b)} > \frac{d}{1-\beta^0(d)} > 0 > \frac{c}{1-\beta^0(c)}$. Note that this is an ordinal assumption on preferences. For example, the first inequality says that each player prefers a constant path in which he defects and the other player cooperates to one in which both players cooperate.

It is useful to illustrate some of the issues discussed in Section 4. First, imagine a restricted ED game in which players choose their actions once and these actions are repeated in all other time periods. This is effectively a one-shot game with payoffs determined by the functions $v_i : A^\infty \to \mathbb{R}, i \in I$. The inequalities insure that the restricted game is prisoners' dilemma. More can be said so long as the stage payoffs are normalized so that the payoffs from (D, D) are zero as in Figure 7.1. Observe that, by IMI, we have $\beta^0(b) < \beta^0(d) < \beta^0(0) < \beta^0(c)$. In turn, this implies that $b > d > 0 > c$. Thus, the stage game as defined by the matrix above is a one-shot prisoners' dilemma game as well. To facilitate interpretation,

In the results below, the following two paths play a recurring role. Let $\mathbf{a}^A$ denote the path $((C, C), (C, C), ...)$ in which the players cooperate in every period. We refer to this path as *intratemporal* cooperation. Let $\mathbf{a}^B$ denote the path $((C, D), (D, C), (C, D)...)$ in which the players alternate between (C, D) and (D, C). Abusing notation, we also use $\mathbf{a}^B$ to denote the alternating sequence that begins with (D, C).

The main results of the section can be summarized as follows. Provided that players are sufficiently patient, there are efficient equilibria. Moreover, in any efficient equilibrium, the play path is eventually equal to either $\mathbf{a}^A$ or $\mathbf{a}^B$. This is the anti-folk result we referred to in the introduction. Which of paths arises in an efficient equilibrium of the game depends on the following inequality. It compares the payoff from alternating between (C, D) and (D, C), as players becomes increasingly patient, with the payoff from cooperating in every period.

$$\lim_{\lambda \to 0} v_i(\mathbf{a}^B) = \frac{b + c}{1 - \beta^0(b) + 1 - \beta^0(c)} \geq \frac{d}{1 - \beta^0(d)} = v_i(\mathbf{a}^A), \quad \forall i \in I. \tag{8}$$

The analysis begins by characterizing the efficient paths when the players have equal weights in the maximization problem in 7. In this case, it is also easy to verify that the corresponding paths are sequentially rational and therefore can be supported in equilibrium.

Theorem 7.1. *If (8) holds, the efficient play path given $\eta = (1, 1)$ is to alternate between (C, D) and (D, C). If (8) does not hold, there exists λ' such that, for any $0 < \lambda < \lambda'$, the efficient play path is the constant path $((C, C), (C, C), ...)$. Moreover, there exists $\overline{\lambda}$ such that for any*

$0 < \lambda < \overline{\lambda}$, *the efficient path given* $\eta = (1,1)$ *can be supported as a subgame perfect equilibrium of the game.*

The next result completes the characterization of the Pareto frontier by considering all directions $\eta \in \mathbb{R}^2_{++}$. It shows that, after a finite number of periods, any efficient path is equal to either a^A or a^B.

Theorem 7.2. *For every* $\lambda \in (0,1], \eta \in \mathbb{R}^2_{++}$ *and* $\mathbf{a} \in P(\lambda,\eta)$, *there exists some time* T *such that the continuation path* ${}_T\mathbf{a}$ *is either* $\mathbf{a}^A$ *or* $\mathbf{a}^B$. *It is* $\mathbf{a}^A$ *if (8) holds and* $\mathbf{a}^B$, *otherwise.*

Along the path $\mathbf{a}^A$ in which both players cooperate, their continuation payoffs at any given point in time are identical. If the path is $\mathbf{a}^B$, the players' continuation payoffs are different but the difference converges to zero as discount factors approach 1. It follows from Theorem 7.2 that, along *any* efficient path the differences in continuation payoffs are eventually negligible. The observation is summarized in the following corollary.

Corollary 7.1. *Given any* $\varepsilon > 0$ *and* $\eta \in \mathbb{R}^2_{++}$, *there exists* $\overline{\lambda}$ *such that for any* $0 < \lambda < \overline{\lambda}$, *there is some* T *such that* $|v_1({}_T\mathbf{a},\lambda) - v_2({}_T\mathbf{a},\lambda)| < \varepsilon$ *for any* $\mathbf{a} \in P(\lambda,\eta)$.

Theorem 7.2 is the main result of this section and the groundwork for our anti-folk result. The remaining difficulty is to check if the efficient paths can be sustained in equilibrium. Once again Theorem 7.2 is helpful. Its proof shows that if an efficient sequence is strictly individually rational at time 0, then it is strictly individually rational at each time t. As a result, in the next theorem we only require the former. To state the result formally, let $IR^{\varepsilon}(\lambda) = \{\mathbf{a} \in A^{\infty} : v_i(\mathbf{a},\lambda) > \varepsilon, i = 1,2\}$

Corollary 7.2. *If there exist* $\varepsilon > 0$ *and* $\eta \in \mathbb{R}^2_{+}$ *such that* $P(\lambda,\eta) \subseteq IR^{\varepsilon}$ *for all* $\lambda \geq 0$, *then there exists* $\overline{\lambda} \in (0,1]$ *such that* $P(\lambda,\eta) \subseteq SPE(\lambda)$ *for all* $0 < \lambda < \overline{\lambda}$, *where* $SPE(\lambda)$ *is the set of all the subgame perfect play paths.*

7.2 Decreasing Marginal Impatience

This section considers arbitrary two-player, symmetric ED games. It shows that, under DMI, any efficient path is eventually constant. There are two possibilities. The first is that all players receive identical payoffs along an efficient path. DMI implies that the players' rates of time preference are identical. This means that there are no gains from intertemporal trade. The other possibility is that one of the players *emerges* as the more patient player and eventually attains his highest feasible payoff. If all efficient paths are of this form, we can no longer guarantee the existence of an efficient equilibrium. Here, the problem is similar to that pointed out in Lehrer and Pauzner [12]. To state the theorem formally, let

$$\overline{A}^i := \operatorname*{argmax}_{a \in A} v_i(a), i \in I, \quad \text{and} \quad A^E := \{a \in A : v_i(a) = v_j(a)\}.$$

Theorem 7.3. *For every $\lambda > 0, \eta \in \mathbb{R}^2_+$, and every $\mathbf{a} \in P(\lambda, \eta)$, there exists some T such that $a^t \in B$ for all $t \geq T$ where $B \in \{\overline{A}^1, \overline{A}^2, A^E\}$.*

In the rest of the section, we consider the implications of Theorem 7.3 for the efficient equilibria of the repeated prisoners' dilemma game. It is clear that paths in which one of the players' continuation payoff is eventually maximized are not sequentially rational for the other player. Therefore, they cannot arise in an equilibrium of the game. The only other potentially efficient paths are the ones in which both players eventually cooperate. From the proof of Theorem 7.3, one can in fact see that, to achieve efficiency, cooperation must start immediately. It remain to verify when this path is efficient. For this, it is sufficient that

$$\frac{d}{1-\beta^0(d)} > \frac{1}{2}\frac{b}{1-\beta^0(b)} + \frac{1}{2}\frac{c}{1-\beta^0(c)}. \tag{9}$$

The inequality says that each player prefers cooperation in every period to a mixed path in which with equal probability he receives his worst or his best stream of outcomes. Note that this requirement is typically imposed on the payoffs of a prisoners' dilemma game. Because cooperation in every period is sequentially rational, we have the following corollary.

Corollary 7.3. *If (9) holds, cooperation in every period is the only path that can arise in an efficient equilibrium of the repeated prisoners' dilemma game. If (9) fails, the game has no efficient equilibrium.*

Combined with the results from the previous section, we reach the notable conclusion that, under *both* increasing and decreasing marginal impatience, the play path in an efficient equilibrium of the PD game, when such exists, is eventually unique. It should be emphasized however that the driving forces behind these results are different. Under

DMI, there is a conflict between efficiency and sequential rationality. As a result, efficiency restricts the paths that can arise in equilibrium. The conflict is similar to that one emphasized by Lehrer and Pauzner [12]. The conflict is less stark in our model because the players are *a priori* symmetric. Consequently, differences in the rate of time preference, which are the trigger of the conflict, can only arise endogenously. This leaves the door open for the players to coordinate on a path along which such differences do not arise. In contrast, under IMI, there is no conflict between efficiency and sequential rationality. In fact, as Theorem 7.1 and its corollaries imply, efficiency promotes sequential rationality by insuring that continuation outcomes are eventually in the mid-segment of the Pareto frontier.

8 Appendix

Let h^t be the complete history observed by all the players at the beginning of time t and let $\sigma^{h_t} \in \Delta(A)$ be the mixed action profile played in that history given a behavioral strategy profile $\sigma \in \Sigma$.

Proof of Lemma 5.1. Fix $i \in I$. First, we show that if players other than i use a constant strategy, then player i's best response is a constant strategy. That is, given any $\alpha_{-i} \in \times_{j \neq i} \Delta A_j$, there exists $\alpha_i \in \Delta A_i$ such that

$$\alpha_i^{con} \in \arg\max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \alpha_{-i}^{con}) \tag{10}$$

To see this, let $\hat{\sigma}_i$ be a best response for player i. Then

$$\begin{aligned} v_i(\hat{\sigma}_i, \alpha_{-i}^{con}) &= \mathbb{E}(g_i(a) + \beta_i(a) v_i(\hat{\sigma}_i^{h^1}, \alpha_{-i}^{con})) \\ &= \mathbb{E} g_i(a) + \mathbb{E}(\beta_i(a) v_i(\hat{\sigma}_i^{h^1}, \alpha_{-i}^{con})) \\ &= g_i(\alpha_i, \alpha_{-i}) + \beta_i(\alpha_i, \alpha_{-i}) v_i(\hat{\sigma}_i, \alpha_{-i}^{con}), \end{aligned} \tag{11}$$

where α_i is the induced mixed action by $\hat{\sigma}_i$ in the first period, and a is in the support of (α_i, α_{-i}). And $\hat{\sigma}_i^{h^1}$ is player i's continuation strategy after history h^1 is realized in the first period. The last equality follows from the fact that $v_i(\hat{\sigma}_i^{h^1}, \alpha_{-i}^{con}) = v_i(\hat{\sigma}_i^{\tilde{h}^1}, \alpha_{-i}^{con}) = v_i(\hat{\sigma}_i, \alpha_{-i}^{con})$, for any possible histories h^1 and $\tilde{h}^1$. Because when the other players use a constant strategy, player i's best response should be independent of histories. From equation (11), we get

$$v_i(\hat{\sigma}_i, \alpha_{-i}^{con}) = \frac{g_i(\alpha_i, \alpha_{-i})}{1 - \beta_i(\alpha_i, \alpha_{-i})} = v_i(\alpha_i^{con}, \alpha_{-i}^{con}),$$

which implies (10).

Similarly, given any $\alpha_i \in \Delta A_i$, there exists $\alpha_{-i} \in \times_{j \neq i} \Delta A_j$ such that

$$\alpha_{-i}^{con} \in \arg \min_{\sigma_{-i} \in \times_{j \neq i} \Sigma_j} v_i(\alpha_i^{con}, \sigma_{-i}). \tag{12}$$

From (10), we have

$$\min_{\sigma_{-i} \in \times_{j \neq i} \Sigma_j} \max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \sigma_{-i}) \leq \min_{\alpha_{-i} \in \times_{j \neq i} \Delta A_j} \max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \alpha_{-i}^{con}) = \min_{\alpha_{-i} \in \times_{j \neq i} \Delta A_j} \max_{\alpha_i \in \Delta A_i} v_i(\alpha_i^{con}, \alpha_{-i}^{con}).$$

For the converse inequality, note that, for every $\sigma_{-i} \in \times_{j \neq i} \Sigma_j$,

$$\max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \sigma_{-i}) \geq \max_{\alpha_i \in \Delta A_i} v_i(\alpha_i^{con}, \sigma_{-i}).$$

Hence,

$$\min_{\sigma_{-i} \in \times_{j \neq i} \Sigma_j} \max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \sigma_{-i}) \geq \min_{\sigma_{-i} \in \times_{j \neq i} \Sigma_j} \max_{\alpha_i \in \Delta A_i} v_i(\alpha_i^{con}, \sigma_{-i}) = \min_{\alpha_{-i} \in \times_{j \neq i} \Delta A_j} \max_{\alpha_i \in \Delta A_i} v_i(\alpha_i^{con}, \alpha_{-i}^{con}).$$

The last equality follows from (12). □

Let $V^{\square} = \prod_{i=1}^{n} [\min_{a \in A} v_i(a), \max_{a \in A} v_i(a)]$. The relation between the sets is illustrated by the following lemma. For simplicity, when we prove Lemma 8.1 and Lemma 6.2, we assume there are only two available pure actions a and b, but all the claims can be generalized to finitely many actions.

Lemma 8.1. $V \subseteq \overline{V}(\lambda) \subseteq V^{\square}$.

Proof of Lemma 8.1. Fix $\lambda \in (0, 1]$. By definition, $V \subseteq \overline{V}(\lambda)$ and $V \subseteq V^{\square}$.

Fix $i \in I$. Suppose $v_i(a) \leq v_i(b)$. It is easy to check that any constant strategy α^{con}, the payoff $v_i(\alpha) = \frac{\alpha(a)g(a) + \alpha(b)g(b)}{1 - \alpha(a)\beta(a) - \alpha(b)\beta(b)}$ is in $[v_i(a), v_i(b)]$.

Note that for any continuation payoff $w \in [v_i(a), v_i(b)]$, we have

$$\lambda g(\alpha) + \beta(\alpha) w \in [v_i(a), v_i(b)], \text{ for any } \alpha. \tag{13}$$

So if the continuation payoff lies in $[v_i(a), v_i(b)]$, no matter how we choose the action today, the payoff will still lie in this interval.

Next we show that the payoff of any strategy is in $[v_i(a), v_i(b)]$. Let w^0 be some feasible continuation payoff and $w^0 + \epsilon \in (v_i(a), v_i(b))$ (ϵ can be any number). By (18), if we take $w + \epsilon$ as continuation payoff, for any current action profile α, the total payoff is $w^1 = \lambda g(\alpha) + \beta(\alpha) w^0 \in [v_i(a), v_i(b)]$. Now take w^1 as the continuation payoff and take any action in the current period. The total payoff w^2 is also in this interval. Repeat this process T times and the limit of w^T as T goes to infinity equals the total payoff if we have started from w^0 at the beginning instead of $w^0 + \epsilon$. Since $w^T \in [v_i(a), v_i(b)]$, the payoff from any strategy lies in $[v_i(a), v_i(b)]$. □

Proof of Lemma 6.1. Fix $i \in I$. Recall that for any $a \in A$, $v_i(a) = \frac{g_i(a)}{1-\beta_i^0(a)}$. *Richness* implies $g_i(a^i) \leq 0 < g_i(\tilde{a}^i)$ and $g_j(a^i) > 0, g_j(\tilde{a}^i) \geq 0$. For any $0 < \epsilon < v_i(\tilde{a}^i)$, there exists α^i whose support is a^i and $\tilde{a}^i$ such that $v_i(\alpha^i) = \frac{g_i(\alpha^i)}{1-\beta_i^0(\alpha^i)} = \epsilon$. We can solve for α^i, which is

$$\alpha^i(a^i) = \frac{g_i(\tilde{a}^i) - (1-\beta_i^0(\tilde{a}^i))\epsilon}{g_i(\tilde{a}^i) - g_i(a^i) + (\beta_i^0(\tilde{a}^i) - \beta_i^0(a^i))\epsilon} \in (0,1) \text{ and } \alpha(\tilde{a}^i) = 1 - \alpha^i(a^i). \tag{14}$$

For any $j \neq i$ and α whose support is a^i and $\tilde{a}^i$, $v_j(\alpha) = \frac{g_j(\alpha)}{1-\beta_j^0(\alpha)} > \epsilon$ if and only if

$$\alpha(a^i)[g_j(a^i) - g_j(\tilde{a}^i) + (\beta_j^0(a^i) - \beta_j^0(\tilde{a}^i))\epsilon] > -g_j(\tilde{a}^i) + (1-\beta_j^0(\tilde{a}^i))\epsilon. \tag{15}$$

It is easy to check that under *richness*, there exists $\overline{\epsilon}^i$ such that for any $0 < \epsilon < \overline{\epsilon}^i$, α^i defined in (14) satisfies inequality (15). Since the number of players is finite, there exists a uniform $\overline{\epsilon} = \min_{i \in I} \overline{\epsilon}^i$ for all $i \in I$ such that the result holds. □

Proof of Theorem 6.1. Take any $\varepsilon > 0$. By Lemma 6.1, for each player $i \in I$, there exists $\overline{\epsilon}$ such that for any $0 < \epsilon < \overline{\epsilon}$, we can find $\alpha^i \in \Delta A$ such that $v_i(\alpha^i) = \epsilon$ and $v_j(\alpha^i) > \epsilon$ for all $j \in I \setminus i$. Take any $0 < \epsilon < \min\{\overline{\epsilon}, \varepsilon\}$. Take any $\mathbf{a} \in A^\infty$ such that $\inf_{t \geq 0, \lambda \in (0,1]} v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$ for all $i \in I$. Let $\overline{g}_i = \max_a g_i(a)$ and $\overline{\beta}_i = \max_a \beta_i(a)$. For each player i, choose an integer μ_i such that $\mu_i > \frac{\overline{g}_i}{\epsilon(1-\beta_i^0(M^i))}$.

Consider the following repeated game strategy for player i:
(A) play a_i^t at period t as long as a^{t-1} was played last period. If player j deviates from (A), then
(B) play M_i^j for μ_j periods and then
(C) play α_i^j thereafter.
If player k deviates in phase (B) or (C), then begin phase (B) again with $j = k$.

If player i deviates in phase (A) and then conforms, he receives at most $\overline{g}_i$ the period he deviates, zero for μ_i periods, and continuation payoff ϵ . His total payoff is no greater than

$$\lambda\overline{g_i} + \overline{\beta}_i[\beta_i(M^i)]^{\mu_i}\epsilon < \varepsilon \leq v_i({}_t\mathbf{a}, \lambda).$$

To see this,

$$\begin{aligned}
&\lambda\overline{g}_i + \overline{\beta}_i[\beta_i(M^i)]^{\mu_i}\epsilon - v_i({}_t\mathbf{a}, \lambda) < \lambda\overline{g}_i + ([1-\lambda(1-\beta_i^0(M^i))]^{\mu_i} - 1)\epsilon < 0 \\
&\Leftrightarrow \frac{\overline{g}_i}{\epsilon} < \frac{1-[1-\lambda(1-\beta_i^0(M^i))]^{\mu_i}}{\lambda(1-\beta_i^0(M^i))}(1-\beta_i^0(M^i)) = \frac{1-\delta^{\mu_i}}{1-\delta}(1-\beta_i^0(M^i)) \\
&\Leftrightarrow \mu_i > \frac{\overline{g}_i}{\epsilon(1-\beta_i^0(M^i))},
\end{aligned}$$

where $\delta = 1-\lambda(1-\beta_i^0(M^i))$ and $\lim_{\lambda\to 0}\frac{1-\delta^{\mu_i}}{1-\delta} = \lim_{\delta\to 1}\frac{1-\delta^{\mu_i}}{1-\delta} = \mu_i$. So the potential gain is less than zero.

If player i deviates in phase (B) when he is being punished, he obtains at most zero the period in which he deviates, and then only lengthens his punishment, postponing the positive continuation payoff ϵ.

If player i deviates in phase (B) when play j is being punished, and then conforms, he receives at most

$$\lambda\overline{g}_i + \overline{\beta}_i[\beta_i(M^i)]^{\mu_i}\epsilon.$$

If he doesn't deviate, he receives at least

$$\frac{g_i(M^j)(1-[\beta_i(M^j)]^{\mu})}{1-\beta_i^0(M^j)} + [\beta_i(M^j)]^{\mu}v_j(\alpha^i),$$

where $1 \leq \mu \leq \mu_j$. Thus the gain to deviating is at most

$$\lambda \overline{g}_i + \overline{\beta}_i [\beta_i(M^i)]^{\mu_i} \epsilon - \frac{g_i(M^j)(1 - [\beta_i(M^j)]^{\mu})}{1 - \beta_i^0(M^j)} - [\beta_i(M^j)]^{\mu} v_j(\alpha^i)$$
$$< \lambda \overline{g}_i + (\epsilon - [\beta_i(M^j)]^{\mu} v_j(\alpha^i)) - \frac{g_i(M^j)(1 - [\beta_i(M^j)]^{\mu})}{1 - \beta_i^0(M^j)}.$$

Since $\overline{g}_i$ and $\frac{g_i(M^j)}{1-\beta_i^0(M^j)}$ are fixed, when λ is close enough to 0, the first part and the last part above approach to zero and the second part is less than zero since $v_j(\alpha^i) > \epsilon$. So the potential gain to deviating is less than zero.

Finally, the argument for why players don't deviate in phase (C) is practically the same as that for phase (A). □

Proof of Lemma 6.2. Fix $\lambda \in (0,1]$. We only need to show for any $v \in \overline{V}(\lambda)$, there exists some $\theta \in [0,1]$ such that $v = \theta v(a) + (1-\theta) v(b)$.

For any a, let $\beta(a) = \beta_i(a)$, for all $i \in I$. Let $\alpha \in \Delta A$. For any $i \in I$,

$$v_i(\alpha) = \frac{\alpha(a) g_i(a) + (1-\alpha(a)) g_i(b)}{1 - [\alpha(a)\beta(a) + (1-\alpha(a))\beta(b)]} = \theta v_i(a) + (1-\theta) v_i(b),$$

where $\theta = \frac{\alpha(a)(1-\beta(a))}{1-[\alpha(a)\beta(a)+(1-\alpha(a))\beta(b)]} \in [0,1]$, which doesn't depend on i.

Let w be a continuation payoff and $w = \theta' v(a) + (1-\theta') v(b)$, where $\theta' \in [0,1]$. For any $\alpha \in \Delta A$, we have

$$\lambda g_i(\alpha) + \beta(\alpha) w_i = [\lambda(1-\beta^0(\alpha))\theta + \beta(\alpha)\theta'] v_i(a) + [\lambda(1-\beta^0(\alpha))(1-\theta) + \beta(\alpha)(1-\theta')] v_i(b).$$

Recall that $\beta(\alpha) = 1 - \lambda(1-\beta^0(\alpha))$. So the payoff is indeed a convex combination of $v_i(a)$ and $v_i(b)$.

Let w^0 be some feasible continuation payoff and $w^0 + \epsilon$ (ϵ can be any number) be a convex combination of $v(a)$ and $v(b)$. Using the same process as in Lemma 8.1, the payoff constructed from w^0 equals the limit payoff constructed from $w^0 + \epsilon$ as T goes to infinity, which is a convex combination of $v(a)$ and $v(b)$.

To prove the second part of the lemma, for any $v = \theta v(a) + (1-\theta) v(b)$, take $\alpha(a) = \frac{\theta(1-\beta^0(b))}{1-((1-\theta)\beta^0(a)+\theta\beta^0(b))} \in [0,1]$. It is straightforward to verify that $v(\alpha) = v$. □

Proof of Theorem 6.2. By Lemma 6.2, for each $v \in V^*$, we can find $\alpha \in \triangle A$ such that $v = v(\alpha)$. Choose $(v'_1, ..., v'_n)$ in the interior of V^* such that $v_i > v'_i$ for all i. Since $v' = (v'_1, ..., v'_n)$ is in the interior of V^* and V^* has full dimension, there exists $\epsilon > 0$ so that for each j,

$$v^j = (v'_1 + \epsilon, ..., v'_{j-1} + \epsilon, v'_j, v'_{j+1} + \epsilon, ..., v'_n + \epsilon) \in V^*.$$

Let T^j be a joint strategy that realizes v^j. Take $\overline{g}_i = \max_a g_i(a)$. For each player i, choose an integer μ_i such that $\mu_i > \frac{\overline{g}_i}{v'_i(1-\beta^0(M^i))}$.

Consider the following repeated game strategy for player i
(A) play α_i each period as long as α was played last period. If player j deviates from (A), then
(B) play M^j_i for μ_j periods and then
(C) play T^j_i thereafter.
If player k deviates in phase (B) or (C), then begin phase (B) again with $j = k$.

The rest of the proof is essentially the same as the proof for Theorem 6.1. □

8.1 Results in Section 7

We begin with some preliminary lemmas. Let the initial pair of weights be $\eta = (\eta_1, \eta_2) \in \mathbb{R}^2_+$. Given any $\mathbf{a} \in P(\lambda, \eta)$, from equation (7), the pair of weights at time t is $\eta^t = (\eta_1 \prod_{\tau=0}^{t-1} \beta_1(a^\tau), \eta_2 \prod_{\tau=0}^{t-1} \beta_2(a^\tau))$.

Lemma 8.2. *Given any $\lambda \in (0, 1]$, $\eta \in \mathbb{R}^2_+$, and $\mathbf{a} \in P(\lambda, \eta)$, we have ${}_t\mathbf{a} \in P(\lambda, \eta^t)$ for each $t > 0$.*

Proof. Suppose there exists $\hat{t}$ such that ${}_{\hat{t}}\mathbf{a} \notin P(\lambda, \eta^{\hat{t}})$, it means there exists $\hat{\mathbf{a}}$ such that

$$\eta_1^{\hat{t}} v_1(\hat{\mathbf{a}}, \lambda) + \eta_2^{\hat{t}} v_2(\hat{\mathbf{a}}, \lambda) > \eta_1^{\hat{t}} v_1({}_{\hat{t}}\mathbf{a}, \lambda) + \eta_2^{\hat{t}} v_2({}_{\hat{t}}\mathbf{a}, \lambda).$$

Construct a path $\mathbf{a}^* = (a^0, ..., a^{\hat{t}-1}, \hat{\mathbf{a}})$. Then

$$
\begin{aligned}
&\eta_1 v_1(\mathbf{a}^*, \lambda) + \eta_2 v_2(\mathbf{a}^*, \lambda) \\
&= \eta_1 \sum_{t=0}^{\hat{t}-1} \prod_{\tau=0}^{t-1} \beta_1(a^\tau) g_1(a^t) + \eta_1^{\hat{t}} v_1(\hat{\mathbf{a}}, \lambda) + \eta_2 \sum_{t=0}^{\hat{t}-1} \prod_{\tau=0}^{t-1} \beta_2(a^\tau) g_2(a^t) + \eta_2^{\hat{t}} v_2(\hat{\mathbf{a}}, \lambda) \\
&> \eta_1 \sum_{t=0}^{\hat{t}-1} \prod_{\tau=0}^{t-1} \beta_1(a^\tau) g_1(a^t) + \eta_1^{\hat{t}} v_1({}_{\hat{t}}\mathbf{a}, \lambda) + \eta_2 \sum_{t=0}^{\hat{t}-1} \prod_{\tau=0}^{t-1} \beta_2(a^\tau) g_2(a^t) + \eta_2^{\hat{t}} v_2({}_{\hat{t}}\mathbf{a}, \lambda) \\
&= \eta_1^t v_1(\mathbf{a}, \lambda) + \eta_2^t v_2(\mathbf{a}, \lambda),
\end{aligned}
$$

which contradicts $\mathbf{a} \in P(\lambda, \eta)$. □

Lemma 8.3. *Given any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_+$, if there exists $a \in A$ such that the constant path $(a, a, ...) \in P(\lambda, \eta)$, then $a \in \overline{A}^1 \cup \overline{A}^2 \cup A^E$.*

Proof. Suppose $(a, a, ...) \in P(\lambda, \eta)$, but $a \notin \overline{A}^1 \cup \overline{A}^2 \cup A^E$, which means $\beta_i(a) > \beta_j(a)$. When t goes to infinity, player i's relative weight will increase to infinity. So repeated play of a can't be efficient. □

Given DMI or IMI, for any $a \in A^E$, $v_1(a) = v_2(a)$. It follows that $\arg\max_{a \in A^E} v_1(a) = \arg\max_{a \in A^E} v_2(a)$. Without loss of generality, assume this set is a singleton and denote it by $\{\overline{a}^e\}$. In Lemma 8.4, 8.5, and 8.6, assume either DMI or IMI.

Lemma 8.4. *Suppose A^E is not empty. Given any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_+$, if there exists $\mathbf{a} \in P(\lambda, \eta)$ and $a^t \in A^E$ for some t, then $a^t = \overline{a}^e$.*

Proof. If A^E is singleton, the result holds trivially. Suppose there exists $\hat{a} \in A^E \setminus \overline{a}^e$ and $a^t = \hat{a}$. Since $\beta_1(\hat{a}) = \beta_2(\hat{a})$, the direction at $t + 1$ is the same as the direction at t. So $\hat{a}$ can still be chosen at $t + 1$. Similarly, $\hat{a}$ can be chosen at any time after t. By construction, the constant path $(\hat{a}, \hat{a}, ...)$ is efficient given η^t. However, since $v_i(\hat{a}) < v_i(\overline{a}^e)$, we have $\eta_1^t v_1(\hat{a}) + \eta_2^t v_2(\hat{a}) < \eta_1^t v_1(\overline{a}^e) + \eta_2^t v_2(\overline{a}^e)$, which means the constant path $(\hat{a}, \hat{a}, ...)$ is strictly dominated by the constant path of $\overline{a}^e$. So $a^t \neq \hat{a}$ for any t. □

Note that given any path $\mathbf{a} \in A^\infty$,

$$v_i({}_t\mathbf{a}, \lambda) = \lambda(1 - \beta_i^0(a^t)) v_i(a^t) + (1 - \lambda(1 - \beta_i^0(a^t))) v_i({}_{t+1}\mathbf{a}, \lambda) \tag{16}$$

Equation (16) says that each $v_i({}_t\mathbf{a}, \lambda)$ is a convex combination of $v_i(a^t)$ and $v_i({}_{t+1}\mathbf{a}, \lambda)$. However, in general, $v({}_t\mathbf{a}, \lambda)$ may not be a convex combination of $v(a^t)$ and $v({}_{t+1}\mathbf{a}, \lambda)$ because players have different discount factors.

Lemma 8.5. *Given any $\lambda \in (0, 1]$, $\eta \in \mathbb{R}^2_+$, and $\mathbf{a} \in P(\lambda, \eta)$, if $a^0 = \overline{a}^e$, then $(\overline{a}^e, \overline{a}^e, ...) \in P(\lambda, \eta)$. Moreover, the Pareto frontier connecting $v({}_1\mathbf{a}, \lambda)$, $v(\mathbf{a}, \lambda)$ and $v(\overline{a}^e)$ is linear, and perpendicular to the direction η.*

Proof. If $\mathbf{a} = (\overline{a}^e, \overline{a}^e, ...)$, this result holds trivially. Suppose $\mathbf{a} \neq (\overline{a}^e, \overline{a}^e, ...)$. Since $\eta^1 = (\eta_1\beta_1(\overline{a}^e), \eta_2\beta_2(\overline{a}^e))$, the direction at $t = 1$ is the same as the direction at $t = 0$. If $\overline{a}^e$ is chosen given η, then $\overline{a}^e$ can also be chosen given η^1. Proceeding like this, we can construct a new path which consists constant play of $\overline{a}^e$, and by construction, this path is efficient. By Lemma 8.2, $(v_1({}_1\mathbf{a}, \lambda), v_2({}_1\mathbf{a}, \lambda))$ is also efficient given η. It implies that $(v_1({}_1\mathbf{a}, \lambda), v_2({}_1\mathbf{a}, \lambda))$, $(v_1(\mathbf{a}, \lambda), v_2(\mathbf{a}, \lambda))$ and $(v_1(\overline{a}^e), v_2(\overline{a}^e))$ are all efficient given η. By equation (16), $v(\mathbf{a}, \lambda)$ is a convex combination of $v({}_1\mathbf{a}, \lambda)$ and $v(\overline{a}^e)$. As a result, $v({}_1\mathbf{a}, \lambda)$, $v(\mathbf{a}, \lambda)$ and $v(\overline{a}^e)$ are on the same linear segment of the Pareto frontier, which is perpendicular to the direction η. □

Lemma 8.6. *Given any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_+$ such that $\frac{\eta_i}{\eta_j} < 1$, for any efficient play path $\mathbf{a} \in P(\lambda, \eta)$ with $v_i(\mathbf{a}, \lambda) < v_j(\mathbf{a}, \lambda)$, if $a^0 = \overline{a}^e$ and ${}_1\mathbf{a} \neq (\overline{a}^e, \overline{a}^e, ...)$, for any $\hat{\eta}$ such that $\frac{\hat{\eta}_i}{\hat{\eta}_j} < \frac{\eta_i}{\eta_j} < 1$ and $\hat{\mathbf{a}} \in P(\lambda, \hat{\eta})$, we have $\hat{a}^0 \neq \overline{a}^e$.*

Proof. Suppose $\hat{a}^0 = \overline{a}^e$. From Lemma 8.5, $v({}_1\mathbf{a}, \lambda)$, $v(\mathbf{a}, \lambda)$ and $v(\overline{a}^e)$ are on the same linear segment of Pareto frontier which is perpendicular to direction η. Similarly, $v(\hat{\mathbf{a}}, \lambda)$ and $v(\overline{a}^e)$ are on the same linear segment of Pareto frontier which is perpendicular to direction $\hat{\eta}$. Since $\frac{\hat{\eta}_i}{\hat{\eta}_j} < \frac{\eta_i}{\eta_j} < 1$, we have $v_i(\hat{\mathbf{a}}, \lambda) \leq v_i(\mathbf{a}, \lambda) < v_j(\mathbf{a}, \lambda) \leq v_j(\hat{\mathbf{a}}, \lambda)$. It is impossible that $(v_i(\overline{a}^e), v_j(\overline{a}^e))$ is on both linear segments of Pareto frontier corresponding to η and $\hat{\eta}$, respectively. □

From now on, we focus on prisoners' dilemma.

Lemma 8.7. *Given any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_{++}$, the constant path $((C, D), (C, D), ...)$ and $((D, C), (D, C), ...)$ can't be efficient.*

Proof. These two cases are symmetric, so we only prove that $((D, C), (D, C), ...)$ can't be efficient. Given any $\eta \in \mathbb{R}^2_{++}$, suppose $((D, C), (D, C)...)$ is efficient. Then there exists some T large enough such that at T, player 1's relative weight $[\frac{\eta_1\beta(b)}{\eta_2\beta(c)}]^T$ is almost 0, while player 2's relative weight $[\frac{\eta_2\beta(c)}{\eta_1\beta(b)}]^T$ is almost infinity. Then at T, choosing (C, D) will improve efficiency, which contradicts the constant path $((D, C), (D, C)...)$ is efficient. □

Lemma 8.8. *Given any $\lambda \in (0, 1]$, $\eta \in \mathbb{R}^2_+$, and $\mathbf{a} \in P(\lambda, \eta)$, if $a^0 = (C, D)$ and $a^1 = (D, C)$, the alternating path $((C, D), (D, C), (C, D), (D, C), ...)$ is efficient. Similarly, if $a^0 = (D, C)$ and $a^1 = (C, D)$, the alternating path $((D, C), (C, D), (D, C), (C, D), ...)$ is efficient.*

Proof. Since the two statements in the lemma are symmetric, we only prove the first one. Since $a^0 = (C, D)$ and $a^1 = (D, C)$, the weight pair at $t = 1$ is $\eta^1 = (\eta_1\beta(c), \eta_2\beta(b))$ and the weight pair at $t = 2$ is $\eta^2 = (\eta_1\beta(c)\beta(b), \eta_2\beta(b)\beta(c))$. Note that the direction η^2 is the same as the direction η. Since $a^0 = (C, D)$, it means given η^2, (C, D) can still be chosen on an efficient path. Similarly, the direction at $t = 3$ is the same as the direction at $t = 1$. Since $a^1 = (D, C)$, then (D, C) can be chosen given η^3. By construction, $((C, D), (D, C), (C, D), (D, C), ...)$ is efficient. □

Proof of Theorem 7.1. Take $\eta = (1, 1)$. By Lemma 8.4, (D, D) will never be chosen on any efficient path because it is strictly dominated by (C, C). So given any efficient play path $\mathbf{a} \in P(\lambda, \eta)$, $a^0 \in \{(C, C), (D, C), (C, D)\}$. Since this is a symmetric game and $\eta_1 = \eta_2$, the case in which $a^0 = (D, C)$ is symmetric with the case in which $a^0 = (C, D)$. So we only consider the following two cases: $a^0 = (C, C)$ and $a^0 = (D, C)$.

Suppose $a^0 = (C, C)$. By Lemma 8.5, the constant path consisting of (C, C) is an efficient path given the assumption that $a^0 = (C, C)$. Denote this constant path by $\mathbf{a}^1$.

Suppose $a^0 = (D, C)$. From Lemma 8.7, the constant path $((D, C), (D, C)...)$ is not efficient. Let t^* be the first time period such that a^{t^*} is either (C, C) or (C, D). Suppose $a^{t^*} = (C, C)$. Take the path $\mathbf{a}^2$ in which $a^t = (D, C)$ for $0 \leq t < t^*$ and $a^t = (C, C)$ for $t \geq t^*$. Applying Lemma 8.5, we know that $\mathbf{a}^2$ is efficient given the assumption that $a^{t^*} = (C, C)$. Suppose $a^{t^*} = (C, D)$. Take the path $\mathbf{a}^3$ in which $a^t = (D, C)$ for $0 \leq t < t^*$, $a^t = (C, D)$ for $t = t^* + 2z$ and $a^t = (D, C)$ for $t = t^* + 2z + 1$, where $z = 0, 1, 2...$. By Lemma 8.8, $\mathbf{a}^3$ is efficient given the assumption that $a^{t^*} = (C, D)$.

Let $s(\mathbf{a},\lambda) = v_1(\mathbf{a},\lambda) + v_2(\mathbf{a},\lambda)$. We get

$$
\begin{aligned}
s(\mathbf{a}^1,\lambda) &= \frac{2d}{1-\beta^0(d)} \\
s(\mathbf{a}^2,\lambda) &= (1-\beta(b)^{t^*})\frac{b}{1-\beta^0(b)} + \beta(b)^{t^*}\frac{d}{1-\beta^0(d)} + (1-\beta(c)^{t^*})\frac{c}{1-\beta^0(c)} + \beta(c)^{t^*}\frac{d}{1-\beta^0(d)} \\
s(\mathbf{a}^3,\lambda) &= (1-\beta(b)^{t^*})\frac{b}{1-\beta^0(b)} + \beta(b)^{t^*}\frac{c+\beta(c)b}{1-\beta(b)\beta(c)} + (1-\beta(c)^{t^*})\frac{c}{1-\beta^0(c)} + \beta(c)^{t^*}\frac{b+\beta(b)c}{1-\beta(b)\beta(c)}.
\end{aligned}
$$

Next we show that $\mathbf{a}^3$ is efficient only if $t^* = 0$ or 1. It is straight forward to check that $s(\mathbf{a}^3,\lambda)$ is the same when t^* is 0 and 1, and strictly decreasing when $t^* \geq 1$. So for $\mathbf{a}^3$ to be efficient, t^* can only be 0 or 1, and its highest value is $s(\mathbf{a}^3,\lambda) = \frac{c+\beta(c)b}{1-\beta(b)\beta(c)} + \frac{b+\beta(b)c}{1-\beta(b)\beta(c)}$. Also $s(\mathbf{a}^3,\lambda)$ is decreasing in λ. Recall that $\beta(a) = 1-\lambda(1-\beta^0(a))$. Take derivative with respect to λ, we get $\frac{ds(\mathbf{a}^3,\lambda)}{d\lambda} < 0$. Intuitively, players benefit from intertemporal trade, and intertemporal trade is made possible because of the difference in discount factors. When $\lambda \to 0$, the difference between discount factors becomes smaller, so is the gain from intertemporal trade. Thus, $s(\mathbf{a}^3,\lambda)$ decreases in λ and

$$\lim_{\lambda\to 0} s(\mathbf{a}^3,\lambda) = \frac{2(b+c)}{1-\beta^0(b)+1-\beta^0(c)}.$$

If equation (8) holds, it is easy to check that $s(\mathbf{a}^3,\lambda) > s(\mathbf{a}^2,\lambda)$ and $s(\mathbf{a}^3,\lambda) > s(\mathbf{a}^1,\lambda)$, which means $\mathbf{a}^3$ is efficient. Otherwise, $s(\mathbf{a}^1,\lambda) > s(\mathbf{a}^2,\lambda)$ and $s(\mathbf{a}^1,\lambda) > s(\mathbf{a}^3,\lambda)$ when $0 < \lambda < \min\{1,\lambda'\}$, where λ' is the solution to $s(\mathbf{a}^1,\lambda) = s(\mathbf{a}^3,\lambda)$ if $s(\mathbf{a}^1,1) < s(\mathbf{a}^3,1)$. In this case, $\mathbf{a}^1$ is efficient.

From the proof we can see that, if (C,C) is chosen on some efficient play path, then (C,C) should be chosen in any period. Conversely, if on some efficient play path, $a^t \neq (C,C)$, then (C,C) will not be chosen in any period. So either $\mathbf{a}^1$ or $\mathbf{a}^3$ is efficient. The "mixed" path $\mathbf{a}^2$ will never be efficient. Moreover, whenever $\eta_1 = \eta_2$, we can choose $a \in \{(D,C),(C,D)\}$, as long as the next period we choose $\{(D,C),(C,D)\} \setminus a$. So we have found all the efficient paths.

Next we show that when λ is small enough, given $\eta = (1,1)$, all the efficient paths can be supported by subgame perfect equilibria. When $\mathbf{a}^1$ is efficient, each player's payoff is $\frac{d}{1-\beta^0(d)} > 0$. When $\mathbf{a}^3$ is efficient, the lowest possible payoff for some player is $\frac{c+\beta(c)b}{1-\beta(b)\beta(c)}$. Since $\lim_{\lambda\to 0}\frac{c+\beta(c)b}{1-\beta(b)\beta(c)} = \frac{b+c}{1-\beta^0(b)+1-\beta^0(c)} > \frac{d}{1-\beta^0(d)} > 0$, there exists λ small enough such that $\frac{c+\beta(c)b}{1-\beta(b)\beta(c)} > 0$. By Theorem 6.1, there exists $\overline{\lambda}$ such that for any $0 < \lambda < \overline{\lambda}$, the efficient paths can be supported as subgame perfect equilibria. □

To prove Theorem 7.2, we need the following lemma.

Lemma 8.9. *Given any $\lambda \in (0,1]$, $\eta \in \mathbb{R}^2_+$ and $\mathbf{a} \in P(\lambda, \eta)$, if $\frac{\eta_1}{\eta_2} < 1$, we have $v_1(\mathbf{a}, \lambda) \leq v_2(\mathbf{a}, \lambda)$ and $a^0 \neq (D, C)$; if $\frac{\eta_1}{\eta_2} > 1$, we have $v_1(\mathbf{a}, \lambda) \geq v_2(\mathbf{a}, \lambda)$ and $a^0 \neq (C, D)$.*

Proof. The two statements in the lemma are symmetric, so we only prove the first one. First we show that if $\frac{\eta_1}{\eta_2} < 1$, we have $v_1(\mathbf{a}, \lambda) \leq v_2(\mathbf{a}, \lambda)$. From Theorem 7.1, when $\hat{\eta}_1 = \hat{\eta}_2$, there exists $\hat{\mathbf{a}} \in P(\lambda, \hat{\eta})$ such that $v_1(\hat{\mathbf{a}}, \lambda) \leq v_2(\hat{\mathbf{a}}, \lambda)$. So if $\frac{\eta_1}{\eta_2} < \frac{\hat{\eta}_1}{\hat{\eta}_2} = 1$, then for any $\mathbf{a} \in P(\lambda, \eta)$, we have $v_1(\mathbf{a}, \lambda) \leq v_1(\hat{\mathbf{a}}, \lambda) \leq v_2(\hat{\mathbf{a}}, \lambda) \leq v_2(\mathbf{a}, \lambda)$.

Suppose $a^0 = (D, C)$. Let T be the first period that $a^t \neq (D, C)$. Such T exists because $v_1(\mathbf{a}, \lambda) \leq v_2(\mathbf{a}, \lambda)$. Since $\beta_1(D, C) < \beta_2(D, C)$, we have $\frac{\eta_1^T}{\eta_2^T} < \frac{\eta_1}{\eta_2} < 1$. We argue that $a^T \neq (C, C)$. Suppose $a^T = (C, C)$. From Lemma 8.5, $v(C, C)$ and $v({}_T\mathbf{a}, \lambda)$ are on the same linear segment of the Pareto frontier. By equation (16) and convexity of the feasible set, this linear segment of Pareto frontier lies below the supporting hyperplane at $v(\mathbf{a}, \lambda)$, which is impossible. Thus, $a^T = (C, D)$. By Lemma 8.8, the play path $\mathbf{a}' = ((D, C), (C, D), (D, C), (C, D), ...) \in P(\lambda, \eta^{T-1})$ and $v_1(\mathbf{a}', \lambda) > v_2(\mathbf{a}', \lambda)$. However, since $\frac{\eta_1^{T-1}}{\eta_2^{T-1}} < \frac{\eta_1}{\eta_2} < 1$, this contradicts the result above. Thus, $a^0 \neq (D, C)$. □

Proof of Theorem 7.2. Take $\lambda \in (0,1]$ and $\eta \in \mathbb{R}^2_{++}$. The case $\eta_1 = \eta_2$ has been proved in Theorem 7.1. Consider the case $\eta_1 \neq \eta_2$. Without loss of generality, assume $0 < \eta_1 < \eta_2$. Take any $\mathbf{a} \in P(\lambda, \eta)$. Since $\eta_1 < \eta_2$, by Lemma 8.9, we have $v_1(\mathbf{a}, \lambda) \leq v_2(\mathbf{a}, \lambda)$. By Lemma 8.4, (D, D) will never be chosen on any efficient path. Also, by Lemma 8.9, $a^0 \neq (D, C)$. Hence, $a^0 \in \{(C, C), (C, D)\}$.

Suppose $a^0 = (C, C)$. By Lemma 8.5, the constant path $((C, C), (C, C), ...)$ is efficient. It implies that (8) doesn't hold. If $\mathbf{a} = ((C, C), (C, C), ...)$, the efficient play path is $\mathbf{a}^A$ from time 0. Next we will show that if $a^0 = (C, C)$, it is impossible that $\mathbf{a} \neq ((C, C), (C, C), ...)$. Suppose $\mathbf{a} \neq ((C, C), (C, C), ...)$ and let T be the first period such that $a^t \neq (C, C)$. For any $0 < t \leq T$, the direction η^t is the same as η. By Lemma 8.9, since $\frac{\eta_1}{\eta_2} < 1$, $a^T \neq (D, C)$. Thus, $a^T = (C, D)$. Since $\beta_1(c) > \beta_2(b)$, we have $\frac{\eta_1^{T+1}}{\eta_2^{T+1}} = \frac{\eta_1 \beta_1(c)}{\eta_2 \beta_2(b)} > \frac{\eta_1}{\eta_2}$. By symmetry, the Pareto frontier corresponding to the direction (η_2, η_1) is a linear segment connecting $(v_1(C, C), v_2(C, C))$ and $(v_2({}_T\mathbf{a}, \lambda), v_1({}_T\mathbf{a}, \lambda))$, where $v_i(C, C) = \frac{d}{1 - \beta^0(d)}$.

If $\frac{\eta_1}{\eta_2} < \frac{\eta_1^{T+1}}{\eta_2^{T+1}} < \frac{\eta_2}{\eta_1}$, then the efficient play path starting from $T+1$ is ${}_{T+1}\mathbf{a} = ((C,C),(C,C),...)$. By Lemma 8.5, $v({}_T\mathbf{a},\lambda)$ and $v(C,C)$ are on the same linear segment of Pareto frontier, and η has the same direction as

$$(v_1(C,C) - v_1({}_T\mathbf{a},\lambda), v_2({}_T\mathbf{a},\lambda) - v_2(C,C)) = \lambda((1-\beta^0(c))\frac{d}{1-\beta^0(d)} - c, b - (1-\beta^0(b))\frac{d}{1-\beta^0(d)}). \tag{17}$$

The path ${}_T\mathbf{a} = ((C,D),(C,C),(C,C),...)$ and the path $\mathbf{a}^A = ((C,C),(C,C),...)$ both are efficient given η, which means they yield the same weighted sum of payoffs, that is,

$$\eta_1(\lambda c + \beta(c)\frac{d}{1-\beta^0(d)}) + \eta_2(\lambda b + \beta(b)\frac{d}{1-\beta^0(d)}) = \eta_1\frac{d}{1-\beta^0(d)} + \eta_2\frac{d}{1-\beta^0(d)}. \tag{18}$$

Equation (17) and (18) hold if and only if (8) holds with equality. But we know in this case, the constant path $\mathbf{a}^A = ((C,C),(C,C),...)$ is not efficient, so we get a contradiction. If $\frac{\eta_1^{T+1}}{\eta_2^{T+1}} \geq \frac{\eta_2}{\eta_1}$, (D,C) can be chosen at $T+1$. By Lemma 8.8, the alternating path $((C,D),(D,C),(C,D),(D,C),...)$ is efficient given η. By Lemma 8.5, $v(\mathbf{a}^B,\lambda)$ and $v(C,C)$ are on the same linear segment of Pareto frontier, and η has the same direction as

$$(v_1(C,C) - v_1(\mathbf{a}^B,\lambda), v_2(\mathbf{a}^B,\lambda) - v_2(C,C)) = (\frac{d}{1-\beta^0(d)} - \frac{\lambda(c+\beta(c)b)}{1-\beta(b)\beta(c)}, \frac{\lambda(b+\beta(b)c)}{1-\beta(b)\beta(c)} - \frac{d}{1-\beta^0(d)}) \tag{19}$$

The path $\mathbf{a}^B = ((C,D),(D,C),(C,D),(D,C),...)$ and the path $\mathbf{a}^A = ((C,C),(C,C),...)$ both are efficient given η, which means they yield the same weighted sum of payoffs, that is,

$$\eta_1\frac{\lambda(c+\beta(c)b)}{1-\beta(b)\beta(c)} + \eta_2\frac{\lambda(b+\beta(b)c)}{1-\beta(b)\beta(c)} = \eta_1\frac{d}{1-\beta^0(d)} + \eta_2\frac{d}{1-\beta^0(d)}. \tag{20}$$

Both (19) and (20) hold if and only if

$$\frac{\lambda(c+\beta(c)b)}{1-\beta(b)\beta(c)} + \frac{\lambda(b+\beta(b)c)}{1-\beta(b)\beta(c)} = \frac{2d}{1-\beta^0(d)}. \tag{21}$$

However, if (21) holds, then we have $\eta_1 = \eta_2$, which contradicts our assumption that $\eta_1 < \eta_2$.

Suppose $a^0 = (C,D)$. By Lemma 8.7, the constant path $((C,D),(C,D),...)$ is not efficient.

Let T be the first period such that $a^t \neq (C, D)$. If $a^T = (C, C)$, from the result above we know that ${}_T\mathbf{a} = \mathbf{a}^A$. If $a^T = (D, C)$, the play path ${}_{T-1}\mathbf{a} = \mathbf{a}^B$ is efficient. This is the unique play path, because by Lemma 8.9, $a^{T+1} \neq (D, C)$, and by the argument above, if $a^{T+1} = (C, C)$, the constant path $\mathbf{a}^A$ will be the unique play path, which contradicts ${}_{T-1}\mathbf{a} = \mathbf{a}^B$ is efficient.

Thus, given any $\lambda \in (0, 1]$, we have constructed all the possible efficient play paths, and in each of them, there exists some T such that ${}_T\mathbf{a}$ is one of $\mathbf{a}^A$ or $\mathbf{a}^B$.

If equation (8) holds, geometrically it means when λ is small enough, the pair of payoffs from constant path $\mathbf{a}^A$ is inside the Pareto frontier. Given any η, $\mathbf{a}^A$ cannot be efficient. So in this case, the continuation path can only be the alternating path $\mathbf{a}^B$. If equation (8) doesn't hold, it means when λ is small enough, the payoffs from alternating path $\mathbf{a}^B$ are inside the Pareto frontier. Therefore, the continuation path can only be the constant path $\mathbf{a}^A$. □

Proof of Corollary 7.1. From 7.2, when λ is small enough, for any $\mathbf{a} \in P(\lambda, \eta)$, the continuation path after some large enough T is either $\mathbf{a}^A$ or $\mathbf{a}^B$. If the continuation path is the constant path $\mathbf{a}^A$, the result trivially holds. If the continuation path is the alternating path $\mathbf{a}^B$, the largest difference between two players' continuation payoffs is

$$\lambda\frac{b + \beta(b)c}{1 - \beta(b)\beta(c)} - \lambda\frac{c + \beta(c)b}{1 - \beta(b)\beta(c)} = \frac{\lambda[b(1 - \beta^0(c)) - c(1 - \beta^0(b))]}{1 - \beta^0(b) + 1 - \beta^0(c) - \lambda(1 - \beta^0(b))(1 - \beta^0(c))},$$

which is increasing in λ. Therefore, given any $\varepsilon > 0$, there exists some $\overline{\lambda}$ such that for any $0 < \lambda < \overline{\lambda}$, the difference between two players' continuation payoffs is less than ε. □

Proof of Corollary 7.2. Take $\varepsilon > 0$ and $\eta \in \mathbb{R}^2_+$ such that for any $\lambda \in (0, 1]$, $P(\lambda, \eta) \subseteq IR^\varepsilon$. We have already proved the case where $\eta_1 = \eta_2$, so here consider the case where $\eta_1 \neq \eta_2$. Without loss of generality, assume $0 < \eta_1 < \eta_2$.

Fix $\lambda \in (0, 1]$. Take any $\mathbf{a} \in P(\lambda, \eta)$. Since $\eta_1 < \eta_2$, by Lemma 8.9, we have $\varepsilon \leq v_1(\mathbf{a}, \lambda) \leq v_2(\mathbf{a}, \lambda)$. We need to show that $\inf_t v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$, for $i = 1, 2$. By Lemma 8.4, (D, D) will never be chosen on any efficient path. Also, by Lemma 8.9, $a^0 \neq (D, C)$. Hence, $a^0 \in \{(C, C), (C, D)\}$.

Suppose $a^0 = (C, C)$. By Lemma 8.5, the constant path $((C, C), (C, C), ...)$ is efficient. If $\mathbf{a} = ((C, C), (C, C), ...)$, then by assumption, this path is strictly individually rational at each t, and we are done. If not, let T be the first period such that $a^t \neq (C,C)$. For any $0 < t \leq T$, the direction η^t is the same as η. By assumption $P(\lambda,\eta) \subseteq IR^{\varepsilon}$, we have $v_i({}_T\mathbf{a},\lambda) \geq \varepsilon$. In Theorem 7.2, we construct all the possible efficient play paths. From the construction, it's easy to check that for each player i, for any t, $v_i({}_t\mathbf{a},\lambda) \geq v_1({}_T\mathbf{a},\lambda) \geq \varepsilon$.

Suppose $a^0 = (C,D)$. Let T be the first period such that $a^t \neq (C,D)$. If $a^T = (C,C)$, then we can continue as the case where $a^0 = (C,C)$. We can show that each player's continuation payoff is bounded below by ε. If $a^T = (D,C)$, by Lemma 8.8, the efficient play path starting from $T-1$ is ${}_{T-1}\mathbf{a} = \mathbf{a}^B$. Since $v_1({}_{T-1}\mathbf{a},\lambda) > v_1(\mathbf{a},\lambda) \geq \varepsilon$, we have $v_2({}_t\mathbf{a},\lambda) \geq v_2({}_T\mathbf{a},\lambda) = v_1({}_{T-1}\mathbf{a},\lambda) > \varepsilon$, for any t. For player 1, $v_1({}_t\mathbf{a},\lambda) \geq v_1(\mathbf{a},\lambda) \geq \varepsilon$, for any t. For player 2, $v_2({}_t\mathbf{a},\lambda) \geq v_2({}_T\mathbf{a},\lambda) > \varepsilon$, for any t. Therefore, each player's continuation payoff is greater than ε.

Thus, given any $\lambda \in (0,1]$, we have constructed all the possible efficient play paths, and shown that in each of them, each player's continuation payoff is above ε. By Theorem 6.1, there exists $\overline{\lambda}$ such that for any $0 < \lambda < \overline{\lambda}$, the efficient play path can be supported by an equilibrium. □

Proof of Theorem 7.3. Take $\lambda \in (0,1]$ and $\eta \in \mathbb{R}^2_+$. Let $\mathbf{a} \in P(\lambda,\eta)$. First we show that there exists some T such that $a^t \in B$ for all $t \geq T$ where $B \in \{\overline{A}^1, \overline{A}^2, A^E\}$.

If there exists $\eta_i = 0$, then $a^t \in \overline{A}^j$ for all t. So in the following proof, assume $\eta \in \mathbb{R}^2_{++}$. For expositional convenience, assume $\overline{A}^1$ and $\overline{A}^2$ are singletons. If $\overline{A}^1 = \overline{A}^2$, it means there is an action profile a^* that yields the highest payoff for both players, then any efficient path is a constant play of a^*. This case is trivial, so we assume $\overline{A}^1 \neq \overline{A}^2$.

Depending on the relative magnitude of players' payoffs, there are two cases to consider. First, $v_i(\mathbf{a},\lambda) = v_j(\mathbf{a},\lambda)$. Suppose $\beta_i(a^0) \neq \beta_j(a^0)$. Without loss of generality, assume $\beta_i(a^0) < \beta_j(a^0)$ and by DMI, $v_i(a^0) < v_j(a^0)$. Note that $v_i(a^0) < v_i(\mathbf{a},\lambda)$, otherwise, $v_j(\mathbf{a},\lambda) = v_i(\mathbf{a},\lambda) \leq v_i(a^0) < v_j(a^0)$, which contradicts $\mathbf{a}$ is efficient. Moreover, $\beta_i(a^0) < \beta_j(a^0)$ implies player i's relative weight decreases at $t = 1$. By Lemma 8.2, $v_i({}_1\mathbf{a},\lambda) \leq v_i(\mathbf{a},\lambda)$. By equation (16), $v_i(\mathbf{a},\lambda)$ is a convex combination of $v_i(a^0)$ and $v_i({}_1\mathbf{a},\lambda)$, but this is impossible, because $v_i(a^0) < v_i(\mathbf{a},\lambda)$ and $v_i({}_1\mathbf{a},\lambda) \leq v_i(\mathbf{a},\lambda)$. So $\beta_i(a^0) = \beta_j(a^0)$ and $v_i(a^0) = v_j(a^0)$. If $v_i(a^0) > v_i(\mathbf{a},\lambda)$, it contradicts the optimality of $\mathbf{a}$; if $v_i(a^0) < v_i(\mathbf{a},\lambda)$, it implies $v_i({}_1\mathbf{a},\lambda) > v_i(\mathbf{a},\lambda)$ and $\mathbf{a}$ is strictly dominated by ${}_1\mathbf{a}$, which contradicts the optimality of $\mathbf{a}$. So we have $v_i(a^0) = v_i(\mathbf{a},\lambda) = v_i({}_1\mathbf{a},\lambda)$. Similarly, applying this argument in each period, we get $\mathbf{a} = (a,a,...)$, where $a \in A^E$.

Second, consider the case $v_i(\mathbf{a}, \lambda) < v_j(\mathbf{a}, \lambda)$. Suppose $a^0 \in A^E$, i.e., $\beta_i(a^0) = \beta_j(a^0)$. By Lemma 8.4, $a^0 = \overline{a}^e$. Since $v_i(\mathbf{a}, \lambda) < v_j(\mathbf{a}, \lambda)$, let T' be the first period such that $\beta_i(a^t) \neq \beta_j(a^t)$. Apply Lemma 8.5 repeatedly, and we can see that for all $0 < t \leq T'$, $(v_i({}_t\mathbf{a}, \lambda), v_j({}_t\mathbf{a}, \lambda))$ are on the same linear segment of Pareto frontier corresponding to direction η. By efficiency, the Pareto frontier is downward sloping, so $v_i({}_t\mathbf{a}, \lambda) < v_i({}_{t-1}\mathbf{a}, \lambda) < v_j({}_{t-1}\mathbf{a}, \lambda) < v_j({}_t\mathbf{a}, \lambda)$, for all $0 < t \leq T'$. Suppose $\beta_i(a^{T'}) > \beta_j(a^{T'})$. By DMI, we have $v_i(a^{T'}) > v_j(a^{T'})$. If $v_j(a^{T'}) \geq v_j(\mathbf{a}, \lambda)$, then $\mathbf{a}$ is strictly Pareto dominated by constant play of $a^{T'}$. So $v_j(a^{T'}) < v_j({}_{T'}\mathbf{a}, \lambda)$. Also because $\beta_i(a^{T'}) > \beta_j(a^{T'})$, player j's relative weight decreases at $T' + 1$. By Lemma 8.2, $v_j({}_{T'+1}\mathbf{a}, \lambda) \leq v_j({}_{T'}\mathbf{a}, \lambda)$, but this contradicts equation (16), because $v_j({}_{T'}\mathbf{a}, \lambda)$ is a convex combination of $v_j(a^{T'})$ and $v_j({}_{T'+1}\mathbf{a}, \lambda)$. Therefore, $\beta_i(a^{T'}) < \beta_j(a^{T'})$. By DMI, it implies $v_i(a^{T'}) < v_j(a^{T'})$. By equation (16), $v_i({}_{T'}\mathbf{a}, \lambda)$ is a convex combination of $v_i(a^{T'})$ and $v_i({}_{T'+1}\mathbf{a}, \lambda)$. Because player i's discount factor is lower, his relative weight decreases. Therefore, we have $v_i({}_{T'+1}\mathbf{a}, \lambda) \leq v_i({}_{T'}\mathbf{a}, \lambda)$. Similarly, player j's relative weight increases, and hence $v_j({}_{T'+1}\mathbf{a}, \lambda) \geq v_j({}_{T'}\mathbf{a}, \lambda)$. By Lemma 8.6, for any $t \geq T'$, we have $a^t \neq \overline{a}^e$. So $\beta_i(a^t) < \beta_j(a^t)$ for each $t \geq T'$. Hence, there exists some T large enough such that player j's relative weight is almost infinity. As a result, for any $t \geq T$, $a^t \in \overline{A}^j$. Applying the same argument as above, when $v_i(\mathbf{a}, \lambda) < v_j(\mathbf{a}, \lambda)$, the other possible case is $\beta_i(a^0) < \beta_j(a^0)$, and there exists some T large enough such that $a^t \in \overline{A}^j$, for all $t \geq T$.

We have shown that along any efficient path $\mathbf{a}$, after some T, the continuation path is a constant path $(a, a, ...)$, where $a \in \overline{A}^1 \cup \overline{A}^2 \cup A^E$. If $g_i(a) < 0$ for some i, it means the continuation payoff for player i is not individually rational. Therefore, $\mathbf{a}$ can't be supported at equilibrium. □

References

[1] D. K. Backus, B. R. Routledge, and S. E. Zin. Exotic preferences for macroeconomists. In *NBER Macroeconomics Annual*, volume 19, pages 319–390. MIT Press, 2004.

[2] G. Becker and C. Mulligan. The endogenous determination of time preference. *The Quarterly Journal of Economics*, 112(3):729–758, 1997.

[3] R. A. Becker. On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households. *The Quarterly Journal of Economics*, 95:375–382, 1980.

[4] P. Dutta. A folk theorem for stochastic games. *Journal of Economic Theory*, 66:1–32, 1995.

[5] L. Epstein. Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, 31(1):133–152, 1983.

[6] L. Epstein. A simple dynamic general equilibrium model. *Journal of Economic Theory*, 41:68–95, 1987.

[7] I. Fisher et al. *The theory of interest*. Macmillan New York, 1930.

[8] M. Friedman. *The optimum quantity of money and other essays*. Aldine, Chicago, 1969.

[9] D. Fudenberg and E. Maskin. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, pages 533–554, 1986.

[10] D. Fudenberg and E. Maskin. On the dispensability of public randomization in discounted repeated games. *Journal of Economic Theory*, 53(2):428–438, 1991.

[11] D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, 1991.

[12] E. Lehrer and A. Pauzner. Repeated games with differential time preferences. *Econometrica*, 67(2):393–412, 1999.

[13] R. Lucas and N. Stokey. Optimal growth with many consumers. *Journal of Economic Theory*, 32(1):139–171, 1984.

[14] H. Uzawa. Time preference, the consumption function, and optimum asset holdings. In J. Wolfe, editor, *Value, Capital and Growth: Papers in Honour of Sir John Hicks*. Aldine, Chicago, 1968.