Repeated Games with Endogenous Discounting*

Asen Kochov Yangwei Song†

December 9, 2014

The 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. The result suggests that the multiplicity of efficient equilibria, traditionally associated with repeated interaction, 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@ur.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 main advantage is analytical 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 ND games for short.

The primary question we ask is: How much of our intuition about repeated interactions relies implicitly on the time-additive specification? Consider the classical prisoners' dilemma game. One implication of the folk theorem for repeated games is that cooperation can be sustained in equilibrium provided that the players are sufficiently patient. A less appealing implication is that the concept of equilibrium loses much of its predictive power. First, there are many equilibria in which the players fail to coordinate on an efficient outcome. There are also many efficient but non-cooperative equilibria. In some of those, one of the players may be held arbitrarily close to his security level throughout the entire game. As we step away from the standard preference specification, the following results emerge. First, the paper confirms the familiar conclusion that every sequentially rational outcome can arise in a subgame perfect equilibrium of the game, as the players become increasingly patient. The potentially surprising finding is that in any equilibrium that attains a first-best outcome, the players must eventually cooperate. In fact, for an intuitive specification of the preferences we consider, cooperation begins immediately. Together, the requirements of efficiency and sequential rationality are thus extremely powerful, selecting cooperation as the unique outcome of the prisoners' dilemma game.

In the literature on endogenous discounting, it is common to assume that discount factors are a strictly monotone function of the underlying payoffs. This is also the assumption under which we investigate the interplay between sequential rationality and efficiency. There are two cases to consider. One is that the marginal impatience of each player i, $1 - \beta_i(a)$, increases the more desirable he finds the constant path $(a, a, ...)$. The polar case of decreasing marginal impatience is defined analogously. The merits of each case have been debated going back to the classical works of Fisher et al. [6, p.72] and Friedman [7, p.30]. Epstein [4, 5] provides a comprehensive summary of the arguments. In this paper, we do not take sides in this debate. The two cases lead to different results. If marginal impatience is decreasing, cooperation begins immediately in any efficient equilibrium. In the polar case of increasing marginal impatience, the uniqueness result is less sharp: cooperation is guaranteed to prevail after some period T. What is interesting about this case, however, is that cooperation may take a different form. Depending on the preference parameters, the players may now cooperate by alternating between their most preferred outcomes in the stage game. In the prisoners' dilemma game, for example, this means that players take turns defecting. We refer to this outcome as **intertemporal cooperation**. Note that, in the standard time-additive model, intertemporal cooperation is never an efficient outcome. 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 dynamics and novel, testable implications.

An axiomatic foundation for the preferences we consider is provided by Epstein [4]. He shows that the utility representation in (1) is implied whenever behavior is stationary and random outcome streams are evaluated according to the expected-utility criterion. Only time separability, arguably the least appealing feature of the standard model, is thus dropped.

2 The Model

Time is discrete and varies over an infinite horizon $t \in \{0, 1, ...\} =: \mathcal{T}$. There is finite set of players $I := \{1, 2, ..., n\}$. In each period t, player i can choose a pure action a_i 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 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 pure action profiles, where $a^t := (a_1^t, ..., a_n^t) \in A$. Given a path $\mathbf{a} = (a^0, a^1, ...,) \in A^\infty$ and a time period $t \in \mathcal{T}$, ${}_t\mathbf{a}$ denotes the continuation path $(a^t, a^{t+1}, ...)$ starting from period t. To describe player i's preferences, define a utility function v_i on A^∞ as follows

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

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 are 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.

An ND game is a tuple $(A, (\beta_i, g_i)_{i \in I})$. By equilibrium, we always mean a subgame perfect equilibrium that induces a deterministic play path. Finally, an equilibrium is efficient if it induces a play path that is not Pareto dominated by any other potentially random play path.

3 Minmax Strategies

One problem in analyzing games with endogenous discounting is that the minmax strategies against each player may change as discount factors are varied. In such instances, it is not possible to fix a sequentially rational strategy profile while independently varying the rate of time preference, as is common in folk theorem analysis. To deal with this problem, in this section we identify a convergence path along which discount factors approach one, yet the minmax strategies against a player remain invariant. A preliminary lemma is needed first. It shows that the search for minmax strategies can be restricted to strategies that are constant. A strategy $\sigma_i, i \in I$, is **constant** if $\alpha_i \in \Delta(A_i)$ is played in every history. Denote each such strategy by α_i^{con}. A profile $\sigma \in \Sigma$ is constant if each σ_i is constant.

Lemma 3.1. *For every player $i \in I$,*

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

From now on, the utility functions are always normalized so that the minmax payoff of each player is zero. It is useful to compute the payoffs from a constant strategy profile $\alpha^{con}, \alpha \in \Delta(A)$. For every $i \in I, \alpha \in \Delta(A)$, let $g_i(\alpha) := \sum_{a \in A} g_i(a)\alpha(a)$, and $\beta_i(\alpha) := \sum_{a \in A} \beta_i(a)\alpha(a)$, where $\alpha(a)$ is the probability assigned to action profile a. Since each constant strategy induces an IID probability measure on A^∞, the ex ante expected payoff from a constant strategy is equal to its expected payoff after any given history. We thus have,

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

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

One can see from (3) that if discount factors are constant, as they are in the standard model, the ranking of constant strategies is completely determined by the stage payoffs $(g_i)_{i \in I}$. Lemma 3.1 can be therefore viewed as an appropriate generalization of the well-known fact that in standard games it is enough to look at minmax strategies in the stage game.

To describe how discount factors converge to one, write each $\beta_i : A \to (0,1), i \in I$, as follows

$$\beta_i(a) = 1 - \lambda(1 - \beta_i^0(a)), \quad \forall a \in A. \tag{5}$$

for some $\lambda \in (0,1]$ and $\beta_i^0 : A \to (0,1)$. This parametrization imposes no restrictions on behavior for a given β_i. If the latter is sufficiently high, it can always be written in this manner. Its purpose is to restrict how discount factors approach one. Specifically, in the rest of the paper, we will be concerned with equilibrium behavior as λ goes to zero. Given the specification in (5), it is also convenient to normalize stage payoffs by taking λg_i instead of g_i. The utilitis v_i over outcome paths can then be written recursively as follows:

$$v_i(\mathbf{a}, \lambda) = \lambda g_i(a^0) + \beta_i(a^0) v_i({}_1\mathbf{a}, \lambda) \tag{6}$$

From now on we use the expressions $v_i(\mathbf{a}, \lambda)$ and $v_i(\mathbf{a})$ interchangeably. The former is preferred when we wish to emphasize that players' preferences change as λ converges to zero.

It remains to verify that the ranking of constant strategies is independent of λ. Plugging (5) in (4) gives:

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

4 The Folk Theorem

This section asks when a given play path $\mathbf{a} \in A^\infty$ can arise in an equilibrium of an ND game. Note that this question differs from the way folk theorems are traditionally stated. In the usual formulation, one is interested whether a feasible, ex ante payoff vector can be attained in equilibrium. There is an important reason why we adopt a formulation that focuses on play paths rather payoffs. When discount factors are constant and the players are symmetric, there are no gains from intertemporal trade: Any feasible payoff can be achieved by constant strategies. With endogenous discounting, the latter is no longer the case. Even if the players are a priori identical, differences in the rate of time preference may emerge endogenously if different players attain different outcomes in the course of the game. The induced heterogeneity creates opportunities for intertemporal trade. As we make the players more and more patient, however, which the folk theorem requires, the gains from intertemporal trade recede. As a result, the feasible set changes as discount factors converge to one and there is no common scale by which to evaluate and compare the achieved payoffs. Note however that the condition is not needed in games with two players.

Another observation is in order. The usual sufficient condition to establish a folk theorem is the full dimensionality assumption introduced by Fudenberg and Maskin [8]. It insures that players can be rewarded for carrying out the punishments against a player who deviates. As was the case for minmax strategies, it is desirable that these strategies are constant. Because of the non-additive nature of the preferences we study, however, we have not been able to find such strategies under full dimensionality. Instead, we propose the following stronger requirement. It is by no means weak: it rules common interest games among things.

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 $v_j(a^i) > 0, v_j(\tilde{a}^i) \geq 0$ for all $j \in I \setminus \{i\}$.

The following lemma demonstrates how we use richness.

Lemma 4.1. *Assume Richness. There exists $\overline{\rho}$ such that for any $0 < \rho < \overline{\rho}$, for any $i \in I$, we can find $\alpha^i \in \Delta A$ such that $v_i(\alpha^i) = \rho$ and $v_j(\alpha^i) > \rho$ for all $j \in I \setminus \{i\}$.*

For any $\varepsilon > 0$ and $\lambda \in (0,1]$, say that a path $\mathbf{a} \in A^\infty$ is **ε-sequentially individually rational**, or **ε-SIR** for short, if $v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$ for all $i \in I, t \in \mathcal{T}$. Let $SIR^\varepsilon(\lambda)$ be the set of all such paths. The next result summarizes our folk theorem.

Theorem 4.1. *Assume Richness. For any $\varepsilon > 0$, there exsists $\overline{\lambda} \in (0,1]$ such that for all $0 < \lambda < \overline{\lambda}$, any path $\mathbf{a} \in SIR^\varepsilon(\lambda)$ can be supported in a subgame perfect equilibrium.*

5 Efficiency

From now on, attention is restricted to symmetric, two-player repeated games. The notion of symmetry needs clarification. As discussed in the previous section, 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. Thus, $\beta_i(a) = f_i(g_i(a))$ for some function f_i. Then, symmetry requires that $g_i = g_j$ and $f_i = f_j$ for all $i, j \in I$. In particular, there is no *a priori* heterogeneity in how discount factors depend on actions: heterogeneity in the rate of time preference can only emerge endogenously when different players attain different outcomes.

It is useful to recall some elementary facts about efficient outcomes. First, given the linearity of expected utility, a play path is efficient if it is not Pareto dominated by a deterministic path. Thus, it is enough to characterize the set of pure efficient paths. Second, since the set of feasible payoffs is convex, every efficient path maximizes a weighted sum of the players' payoffs. In particular, let $P(\lambda, \eta)$ be the set of efficient paths that solve the maximization problem

$$\max_{\mathbf{a} \in A^\infty} \eta_1 v_1(\mathbf{a}, \lambda) + \eta_2 v_2(\mathbf{a}, \lambda), \tag{7}$$

given the 'Pareto weights' $\eta := (\eta_1, \eta_2) \in \mathbb{R}^2_+$. Also, let $P(\lambda) = \cup_{\eta \in \mathbb{R}_{++}} P(\lambda, \eta)$. As was explained in the introduction, the next two sections characterize the set of efficient paths under the assumptions of increasing and decreasing marginal impatience respectively. These can be formalized as follows.

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

The assumption of **decreasing marginal impatience**, or **DMI** for short, is defined analogously.

5.1 Increasing Marginal Impatience

Assume that marginal impatience is increasing. The focus of this section is the repeated prisoners' dilemma game. Let the action space A and the stage payoffs $(g_i)_{i \in I}$ be as in Figure 1, where, as usual, C stands for the action 'cooperate' and D for 'defect'.

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

Figure 1: The prisoners' dilemma

Because discount factors depend on stage payoffs only and this dependence is identical across players, we can write $\beta^0(d) := \beta^0_1(D, C) = \beta^0_2(C, D)$. Similarly for all other outcomes. Assume that $\frac{d}{1-\beta^0(d)} > \frac{c}{1-\beta^0(c)} > 0 > \frac{b}{1-\beta^0(b)}$. 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. As is typical, we also impose the following requirement on the payoffs of a prisoners' dilemma game.

$$\frac{d}{1-\beta^0(d)} > \frac{1}{2}\frac{b}{1-\beta^0(b)} + \frac{1}{2}\frac{c}{1-\beta^0(c)}. \tag{8}$$

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.

We now formalize the two forms of cooperation we consider. Let $\mathbf{a}^C$ denote the constant path $((C, C), (C, C), ...)$ in which the players cooperate in every period. We refer to this path as **intratemporal cooperation**. Let $\mathbf{a}^A$ denote the path $((C, D), (D, C), (C, D), (D, C)...)$

in which the players alternate between (C, D) and (D, C). We refer to this path as **intertemporal cooperation**. Also, let $\mathcal{A}_1^C$ be the set of all pure paths $\mathbf{a}$ such that (D, C) is played until some period T and ${}_T\mathbf{a} = \mathbf{a}^c$. These are the paths in which player 1 attains his highest stage payoff up until some period T, after which the players cooperate. Define $\mathcal{A}_2^C$ analogously and let $\mathcal{A}^C := \mathcal{A}_1^C \cup \mathcal{A}_1^C$. Finally, define $\mathcal{A}_1^A, \mathcal{A}_2^A$, and $\mathcal{A}^A$ to be the paths in which cooperation is intertemporal. The next theorem is the main result in this section.

Theorem 5.1. *Fix $\lambda \in (0, 1]$. If $v_1(\mathbf{a}^C, \lambda) + v_2(\mathbf{a}^C, \lambda) > v_1(\mathbf{a}^A, \lambda) + v_2(\mathbf{a}^A, \lambda)$, then $P(\lambda) = \mathcal{A}^C$ and the unique efficient path given the Pareto weights $\eta = (1, 1)$ is $\mathbf{a}^C$. If $v_1(\mathbf{a}^C, \lambda) + v_2(\mathbf{a}^C, \lambda) < v_1(\mathbf{a}^A, \lambda) + v_2(\mathbf{a}^A, \lambda)$, then $P(\lambda) = \mathcal{A}^A$ and $\mathbf{a}^A$ is the unique efficient path given $\eta = (1, 1)$.*

Having characterized the efficient paths, it remains to apply Theorem 4.1 to deduce that any such sequentially rational path can indeed be sustained in a subgame perfect equilibrium, provided that the players are sufficiently patient. Remarkably, the structure of efficient paths implies that individual rationality implies sequential rationality. We summarize this observation in the next corollary. First, for any $\varepsilon > 0$, let $IR^\varepsilon(\lambda) = \{\mathbf{a} \in A^\infty : v_i(\mathbf{a}, \lambda) \geq \varepsilon, \forall i \in I\}$.

Corollary 5.1. *Take any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_+$. For any $\varepsilon > 0$, if path $\mathbf{a} \in P(\lambda, \eta) \cap IR^\varepsilon(\lambda)$, then $\mathbf{a} \in SIR^\varepsilon(\lambda)$.*

Corollary 5.2. *For any $\varepsilon > 0$, there exists $\overline{\lambda} \in (0, 1]$ such that, for any $0 < \lambda < \overline{\lambda}$, we can find $0 < \underline{\eta} < 1$ such that for any $\eta \in \mathbb{R}^2_+$ and $\underline{\eta} < \frac{\eta_1}{\eta_2} < \frac{1}{\underline{\eta}}$, the play path $\mathbf{a} \in P(\lambda, \eta)$ can be supported in an equilibrium of the game.*

5.2 Decreasing Marginal Impatience

This section considers arbitrary two-player, symmetric ND 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 5.2. *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 $B \in \{\overline{A}^1, \overline{A}^2, A^E\}$.*

Now specialize Theorem 5.2 to 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 5.2, one can see that, to achieve efficiency, cooperation must in fact start immediately. Thus, we have the following corollary.

Corollary 5.3. *Intratemporal cooperation is the only path that can arise in an efficient equilibrium of the repeated prisoners' dilemma game.*

6 Appendix

Writing $v =: (v_1, ..., v_n)$ for the profile of utility functions, we now define the following sets which will be used in the rest of the paper

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

where 'co' denotes the convex hull of a set. Thus, $\overline{V}(\lambda)$ is the set of feasible payoff vectors in the repeated game and V is the convex hull of all payoff vectors attainable by constant pure strategies. Note that the latter is independent of λ but not the former. Unlike standard games, V may also be a strict subset of $\overline{V}(\lambda)$ for every λ. This is the case whenever there are gains from what we call intertemporal cooperation. Finally, define the corresponding subsets of strictly individually-rational payoffs

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

Let h^t be the complete history observed by all the players at the beginning of time t.

Proof of Lemma 3.1. 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 $\alpha_{-i} \in \times_{k \neq i} \Delta A_k$, there exists $\alpha_i \in \Delta A_i$ such that

$$\alpha_i^{con} \in \arg\max_{\sigma_i \in \Sigma_i} v_i(\sigma_i, \alpha_{-i}^{\text{con}}) \tag{9}$$

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

$$\begin{aligned} v_i(\hat{\sigma}_i, \alpha_{-i}^{\text{con}}) &= \mathbb{E}[g_i(a) + \beta_i(a) v_i(\hat{\sigma}_i^{h^1}, \alpha_{-i}^{\text{con}})] \\ &= \mathbb{E} g_i(a) + \mathbb{E}[\beta_i(a) v_i(\hat{\sigma}_i^{h^1}, \alpha_{-i}^{\text{con}})] \\ &= g_i(\alpha_i, \alpha_{-i}) + \beta_i(\alpha_i, \alpha_{-i}) v_i(\hat{\sigma}_i, \alpha_{-i}^{\text{con}}), \end{aligned} \tag{10}$$

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 $v_i(\hat{\sigma}_i^{h^1}, \alpha_{-i}^{\text{con}}) = v_i(\hat{\sigma}_i^{\tilde{h}^1}, \alpha_{-i}^{\text{con}}) = v_i(\hat{\sigma}_i, \alpha_{-i}^{\text{con}})$, 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 (10), we get

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

which implies (9).

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

$$\alpha_{-i}^{\text{con}} \in \arg\min_{\sigma_{-i} \in \times_{k \neq i} \Sigma_k} v_i(\alpha_i^{con}, \sigma_{-i}). \tag{11}$$

From (9), we have

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

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

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

Hence,

$$\min_{\sigma_{-i}\in\times_{k\neq i}\Sigma_k}\max_{\sigma_i\in\Sigma_i} v_i(\sigma_i,\sigma_{-i}) \geq \min_{\sigma_{-i}\in\times_{k\neq i}\Sigma_k}\max_{\alpha_i\in\Delta A_i} v_i(\alpha_i^{con},\sigma_{-i}) = \min_{\alpha_{-i}\in\times_{k\neq i}\Delta A_k}\max_{\alpha_i\in\Delta A_i} v_i(\alpha_i^{con},\alpha_{-i}^{\text{con}}). \tag{12}$$

The last equality follows from (11). □

For each $i \in I$, let M^i be the minmax strategy against player i. Normalize the minmax payoff for each player to zero, i.e., $g_i(M^i) = 0$.

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.

Lemma 6.1. $V \subseteq \overline{V}(\lambda) \subseteq V^\square$.

Proof of Lemma 6.1. Fix $\lambda \in (0,1]$. By definition, $V \subseteq \overline{V}(\lambda)$ and $V \subseteq V^\square$. Next we show $\overline{V}(\lambda) \subseteq V^\square$. We prove this result by assuming only two pure actions a and a' are available. This result is easily generalized to finitely many actions. Fix $i \in I$. Suppose $v_i(a) \leq v_i(a')$. It is easy to check that, for any constant strategy α^{con}, the payoff $v_i(\alpha) = \frac{\alpha(a)g(a)+\alpha(a')g(a')}{1-\alpha(a)\beta(a)-\alpha(a')\beta(a')}$ is in $[v_i(a), v_i(a')]$.

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

$$\lambda g_i(\alpha) + \beta_i(\alpha)w \in [v_i(a), v_i(a')], \text{ for any } \alpha \in \Delta A. \tag{13}$$

So if the continuation payoff lies in $[v_i(a), v_i(a')]$, 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 $[v_i(a), v_i(a')]$. Let w^0 be some feasible continuation payoff and $w^0 + \rho \in \big(v_i(a), v_i(a')\big)$ (ρ can be any number). By (13), if we take $w + \rho$ as continuation payoff, for any current action profile $\alpha^1 \in \Delta A$, the total payoff is $w^1 = \lambda g_i(\alpha^1) + \beta_i(\alpha^1)w^0 \in [v_i(a), v_i(a')]$. Now take w^1 as the continuation payoff and take any action in the current period. The total payoff is $w^2 = \lambda g_i(\alpha^2) + \beta_i(\alpha^2)w^1 \in [v_i(a), v_i(a')]$. 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 $w^0 + \rho$. Since $w^T \in [v_i(a), v_i(a')]$, any feasible payoff lies in $[v_i(a), v_i(a')]$. □

Proof of Lemma 4.1. 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 that there exist action profiles a^i, $\tilde{a}^i \in A$ such that $g_i(a^i) \leq 0 < g_i(\tilde{a}^i)$ and $g_j(a^i) > 0, g_j(\tilde{a}^i) \geq 0$ for all $j \in I \setminus \{i\}$. For any $0 < \rho < v_i(\tilde{a}^i)$, let

$$\kappa^i(a^i) := \frac{g_i(\tilde{a}^i) - (1 - \beta_i^0(\tilde{a}^i))\rho}{g_i(\tilde{a}^i) - g_i(a^i) + \big(\beta_i^0(\tilde{a}^i) - \beta_i^0(a^i)\big)\rho} \in (0,1) \text{ and } \kappa(\tilde{a}^i) := 1 - \kappa^i(a^i).$$

Direct verification shows that $v_i(\kappa^i) = \frac{g_i(\kappa^i)}{1-\beta_i^0(\kappa^i)} = \rho$. Moreover, there exists $\overline{\rho}^i$ such that for any $0 < \rho < \overline{\rho}^i$, we have $v_j(\kappa^i) = \frac{g_j(\kappa^i)}{1-\beta_j^0(\kappa^i)} > \rho$, for all $j \neq i$.

Since the number of players is finite, $\overline{\rho} = \min_{i \in I} \overline{\rho}^i$ exists. This completes the proof. □

In the proof for Theorem 4.1, the subgame perfect equilibrium strategies are the same as the ones constructed in Fudenberg and Maskin [8]. The difficulty is the existence of the strategies that incentivize the other players to punish the deviator. This problem is solved by the assumption richness. The proof is constructed by checking that there is no profitable one-shot deviation at each phase of the strategy. By Theorem 4.2 in Fudenberg and Tirole [11], the one-shot deviation principle is valid here because endogenous discounting games are continuous at infinity.[1]

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

$$v_i({}_t\mathbf{a}, \lambda) = \lambda\big(1 - \beta_i^0(a^t)\big)v_i(a^t) + \Big(1 - \lambda\big(1 - \beta_i^0(a^t)\big)\Big)v_i({}_{t+1}\mathbf{a}, \lambda) \tag{14}$$

Equation (14) 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 $\beta_1^0(a^t) \neq \beta_2^0(a^t)$.

Lemma 6.2. *Given $\lambda \in (0,1]$, for any path $\mathbf{a} \in A^\infty$ such that there exists $\varepsilon > 0$ and $\mathbf{a} \in SIR^\varepsilon(\lambda)$ if and only if there exists $\alpha^i \in \Delta(A)$ such that $v_i({}_t\mathbf{a}, \lambda) \geq v_i(\alpha^i) > 0$ for all $t \in \mathcal{T}$.*

Proof. By definition, if $\mathbf{a} \in A^\infty$ is uniformly sequentially rational, there exists $\alpha^i \in \Delta A$ such that $v_i({}_t\mathbf{a}, \lambda) \geq v_i(\alpha^i) > 0$, for all $i \in I$, $t \in \mathcal{T}$ and $\lambda \in (0,1]$.

"Only if" part: Fix $i \in I$ and $\lambda \in (0,1]$. By equation (14), $v_i(\mathbf{a}, \lambda)$ is a convex combination of $v_i(a^t)$. Since $v_i(\mathbf{a}, \lambda) \geq \varepsilon$, there exists $\tilde{a}^i \in A$ such that $v_i(\tilde{a}^i) \geq \varepsilon$. For any $0 < \rho \leq \varepsilon$, let

[1]Endogenous discounting games are continuous at infinity because the discount factors are less than one and stage payoffs are bounded.

$p^i := \frac{g_i(\tilde{a}^i) - \left(1-\beta_i^0(\tilde{a}^i)\right)\rho}{g_i(\tilde{a}^i) + \left(\beta_i^0(\tilde{a}^i) - \beta_i^0(M^i)\right)\rho} \in (0,1)$. Let $\alpha^i \in \Delta A$ be the strategy such that for each action $a \in A$, α^i assigns a with probability $p^i M^i(a) + (1-p^i)\tilde{a}^i(a)$. Direct verification shows that $0 < v_i(\alpha^i) = \rho \leq \varepsilon$. Thus, $v_i({}_t\mathbf{a}, \lambda) \geq v_i(\alpha^i) > 0$.

"If" part: Let $\varepsilon^i := v_i(\alpha^i)$ and $\varepsilon := \min_{i\in I} \varepsilon^i$. Thus, $v_i({}_t\mathbf{a}, \lambda) \geq v_i(\alpha^i) = \varepsilon^i \geq \varepsilon > 0$. □

Proof of Theorem 4.1. Take $\varepsilon > 0$. By Lemma 4.1, there exists $\overline{\rho}$ such that for any $0 < \rho < \overline{\rho}$, for each player $i \in I$, we can find $\alpha^i \in \Delta A$ such that $v_i(\alpha^i) = \rho$ and $v_j(\alpha^i) > \rho$ for all $j \in I \setminus i$. Take any $0 < \rho < \min\{\overline{\rho}, \varepsilon\}$. Let $\kappa^i \in \Delta A$ be the mixed action profile such that $v_i(\kappa^i) = \rho$ and $v_j(\kappa^i) > \rho$ for all $j \in I \setminus i$. Let $\overline{g}_i = \max_a g_i(a)$ and $\overline{\beta}_i = 1 - \lambda\big(1 - \max_a \beta_i^0(a)\big)$. For each player i, choose an integer μ_i such that $\mu_i > \frac{\overline{g}_i}{v_i(\kappa^i)\left(1-\beta_i^0(M^i)\right)}$. Let $\delta := 1 - \lambda\big(1 - \beta_i^0(M^i)\big)$. For λ small enough, the following inequality holds,

$$\frac{\overline{g}_i}{v_i(\kappa^i)\big(1 - \beta_i^0(M^i)\big)} < \frac{1-\delta^{\mu_i}}{1-\delta}, \tag{15}$$

since $\lim_{\lambda\to 0} \frac{1-\delta^{\mu_i}}{1-\delta} = \lim_{\delta \to 1} \frac{1-\delta^{\mu_i}}{1-\delta} = \mu_i$ and the assumption that $\mu_i > \frac{\overline{g}_i}{v_i(\kappa^i)\left(1-\beta_i^0(M^i)\right)}$. By continuity, there exists $\overline{\lambda}' \in (0,1]$ such that for any $\lambda \in (0, \overline{\lambda}')$ inequality (15) holds. Moreover, when λ is close to zero, we have the following inequality,

$$\lambda \overline{g}_i + \big(v_i(\kappa^i) - [\beta_i(M^j)]^{\mu} v_i(\kappa^j)\big) - \frac{g_i(M^j)\big(1 - [\beta_i(M^j)]^{\mu}\big)}{1 - \beta_i^0(M^j)} < 0, \tag{16}$$

where $1 \leq \mu \leq \mu_j$. 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 term and the last term above approach zero. The second term is less than zero since $v_i(\kappa^j) > v_i(\kappa^i)$. By continuity, there exists $\overline{\lambda}'' \in (0,1]$ such that for any $\lambda \in (0, \overline{\lambda}'')$ inequality (16) holds. Let $\overline{\lambda} := \min\{\overline{\lambda}', \overline{\lambda}''\}$. Take any $\lambda \in (0, \overline{\lambda})$ and $\mathbf{a} \in SIR^{\varepsilon}(\lambda)$. By definition, we have $v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$, for all $i \in I$ and $t \in \mathcal{T}$.

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 κ_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 $v_i(\kappa^i)$. His total payoff is no greater than $\lambda\overline{g}_i + \overline{\beta}_i[\beta_i(M^i)]^{\mu_i}v_i(\kappa^i)$. The gain from deviating is less than

$$\lambda\overline{g}_i + \overline{\beta}_i[\beta_i(M^i)]^{\mu_i}v_i(\kappa^i) - v_i({}_t\mathbf{a}, \lambda),$$

which is less than

$$\lambda\overline{g}_i + \Big([1-\lambda\big(1-\beta_i^0(M^i)\big)]^{\mu_i} - 1\Big)v_i(\kappa^i) \tag{17}$$

because $\overline{\beta}_i < 1$ and $v_i(\kappa^i) < \varepsilon \le v_i({}_t\mathbf{a}, \lambda)$. Direct verification shows that (17) is less than 0 if and only if (15) is satisfied. Sicne $\lambda < \overline{\lambda}'$, 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 $v_i(\kappa^i)$. If player i deviates in phase (B) when play j is being punished, and then conforms, he receives at most $\lambda\overline{g}_i + \overline{\beta}_i[\beta_i(M^i)]^{\mu_i}v_i(\kappa^i)$. If he doesn't deviate, he receives at least $\frac{g_i(M^j)\big(1-[\beta_i(M^j)]^{\mu}\big)}{1-\beta_i^0(M^j)} + [\beta_i(M^j)]^{\mu}v_i(\kappa^j)$, where $1 \le \mu \le \mu_j$. Thus the gain to deviating is at most

$$\lambda\overline{g}_i + \big(v_i(\kappa^i) - [\beta_i(M^j)]^{\mu}v_i(\kappa^j)\big) - \frac{g_i(M^j)\big(1-[\beta_i(M^j)]^{\mu}\big)}{1-\beta_i^0(M^j)}.$$

Since $\lambda < \overline{\lambda}''$, 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). Therefore, given any $\varepsilon > 0$, there exists an upper bound $\overline{\lambda}$ such that for any $0 < \lambda < \overline{\lambda}$, any path $\mathbf{a} \in SIR^{\varepsilon}(\lambda)$ can be supported in equilibrium. □

Proof of Lemma **??**. Since $\beta_i = \beta_j$ for all $i, j \in I$, the subscript is omitted and denote the discount factor by β. Assume only two pure actions a and a' are available. This result is easily generalized to finitely many actions. 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 $v = \theta v(a) + (1-\theta)v(a')$.

First we show that the payoff from any constant mixed strategy is a convex combination of payoffs from constant pure strategies. Take $\alpha \in \Delta A$ and let $\theta := \frac{\alpha(a)\big(1-\beta(a)\big)}{1-\Big(\alpha(a)\beta(a)+\big(1-\alpha(a)\big)\beta(a')\Big)} \in$

$[0,1]$. For any $i \in I$, by direct verification, we have

$$v_i(\alpha) = \frac{\alpha(a)g_i(a) + \big(1-\alpha(a)\big)g_i(a')}{1 - \Big(\alpha(a)\beta(a) + \big(1-\alpha(a)\big)\beta(a')\Big)} = \theta v_i(a) + (1-\theta)v_i(a').$$

Next we show that the payoff from any non-constant strategy is a convex combination of payoffs from constant pure strategies. Let w be a vector of continuation payoffs and $w = \theta' v(a) + (1-\theta')v(a')$, where $\theta' \in [0,1]$. For any $\alpha \in \Delta A$, we have

$$\begin{aligned}\lambda g_i(\alpha) + \beta(\alpha)w_i = &\Big(\lambda\big(1-\beta^0(\alpha)\big)\theta + \beta(\alpha)\theta'\Big)v_i(a) \\ &+ \Big(\lambda\big(1-\beta^0(\alpha)\big)(1-\theta) + \beta(\alpha)(1-\theta')\Big)v_i(a').\end{aligned}$$

Recall that $\beta(\alpha) = 1 - \lambda\big(1-\beta^0(\alpha)\big)$. Hence $\lambda g_i(\alpha) + \beta(\alpha)w_i$ is indeed a convex combination of $v_i(a)$ and $v_i(a')$.

Let w^0 be some vector of feasible continuation payoffs and $w^0 + \rho$ (ρ can be any number) be a convex combination of $v(a)$ and $v(a')$. Using the same process as in Lemma 6.1, the vector of payoffs constructed from w^0 equals the vector of limit payoffs constructed from $w^0 + v_i(\kappa^i)$ as T goes to infinity, which is a convex combination of $v(a)$ and $v(a')$.

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

Proof of Theorem **??**. By Lemma **??**, 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 $\zeta > 0$ so that for each j,

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

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 $\mu_i > \frac{\overline{g}_i}{v'_i\big(1-\beta^0(M^i)\big)}$.

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_i^j for μ_j periods and then
(C) play T_i^j 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 4.1. □

6.1 Results in Section 5

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 $\eta^t(\mathbf{a}) = \left(\eta_1 \prod_{\tau=0}^{t-1} \beta_1(a^\tau), \eta_2 \prod_{\tau=0}^{t-1} \beta_2(a^\tau)\right)$. When the efficient play path $\mathbf{a}$ is clear in the context, the argument is omitted. The following lemma is used repeatedly. The proof is straightforward, and hence omitted.

Lemma 6.3. *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 all $t > 0$.*

In Lemma 6.4, 6.5, and 6.6, assume either DMI or IMI. By DMI or IMI, for any $a \in A^E$, $v_1(a) = v_2(a)$. Without loss of generality, assume the set $\arg\max_{a \in A^E} v_i(a)$ is a singleton and denote it by $\{\overline{a}^e\}$.

Lemma 6.4. *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 $\overline{a}^e$. It contradicts $(\hat{a}, \hat{a}, , ...)$ is efficient. Hence, $\hat{a}$ cannot be chosen on any efficient path, that is, if $a^t \in A^E$, then $a^t = \overline{a}^e$. □

Lemma 6.5. *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 $\mathbf{a} = (\bar{a}^e, \bar{a}^e, ...)$, this result holds trivially. Suppose $\mathbf{a} \neq (\bar{a}^e, \bar{a}^e, ...)$. Since $\eta^1 = \big(\eta_1\beta_1(\bar{a}^e), \eta_2\beta_2(\bar{a}^e)\big)$, the direction at $t = 1$ is the same as the direction at $t = 0$. If $\bar{a}^e$ is chosen given η, then $\bar{a}^e$ can also be chosen given η^1. Proceeding like this, we can construct a new path which consists constant play of $\bar{a}^e$, and by construction, this path is efficient. By Lemma 6.3, $\big(v_1({}_1\mathbf{a}, \lambda), v_2({}_1\mathbf{a}, \lambda)\big)$ is also efficient given η. It implies that $\big(v_1({}_1\mathbf{a}, \lambda), v_2({}_1\mathbf{a}, \lambda)\big)$, $\big(v_1(\mathbf{a}, \lambda), v_2(\mathbf{a}, \lambda)\big)$ and $\big(v_1(\bar{a}^e), v_2(\bar{a}^e)\big)$ are all efficient given η. As a result, $v({}_1\mathbf{a}, \lambda)$, $v(\mathbf{a}, \lambda)$ and $v(\bar{a}^e)$ are on the same linear segment of the Pareto frontier, which is perpendicular to the direction η. □

Lemma 6.6. *Given any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_+$ such that for any efficient play path $\mathbf{a} \in P(\lambda, \eta)$ with $v_i(\mathbf{a}, \lambda) < v_j(\mathbf{a}, \lambda)$, if $a^0 = \bar{a}^e$, for any $\hat{\eta}$ such that $\frac{\hat{\eta}_i}{\hat{\eta}_j} < \frac{\eta_i}{\eta_j}$ and $\hat{\mathbf{a}} \in P(\lambda, \hat{\eta})$, we have $\hat{a}^0 \neq \bar{a}^e$.*

Proof. Suppose $\hat{a}^0 = \bar{a}^e$. From Lemma 6.5, $v(\mathbf{a}, \lambda)$ and $v(\bar{a}^e)$ are on the same linear segment of Pareto frontier which is perpendicular to direction η. Similarly, $v(\hat{\mathbf{a}}, \lambda)$ and $v(\bar{a}^e)$ are on the same linear segment of Pareto frontier which is perpendicular to direction $\hat{\eta}$. Since $\frac{\hat{\eta}_i}{\hat{\eta}_j} < \frac{\eta_i}{\eta_j}$, 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 $\big(v_i(\bar{a}^e), v_j(\bar{a}^e)\big)$ is on both linear segments of Pareto frontier corresponding to η and $\hat{\eta}$, respectively. □

From now on, assume IMI and focus on prisoners' dilemma.

Lemma 6.7. *Given any $\lambda \in (0, 1]$ and $\eta \in \mathbb{R}^2_{++}$, the constant path $\big((C, D), (C, D), ...\big)$ and $\big((D, C), (D, C), ...\big)$ cannot be efficient.*

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

Lemma 6.8. *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 $\big((C, D), (D, C), (C, D), (D, C), ...\big)$ is efficient. Similarly, if $a^0 = (D, C)$ and $a^1 = (C, D)$, the alternating path $\big((D, C), (C, D), (D, C), (C, D), ...\big)$ 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 $\eta^1 = \big(\eta_1\beta(b), \eta_2\beta(d)\big)$

and the weight pair at $t = 2$ is $\eta^2 = \big(\eta_1\beta(b)\beta(d), \eta_2\beta(d)\beta(b)\big)$. 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, $\big((C, D), (D, C), (C, D), (D, C), ...\big)$ is efficient. □

Let $v_1(C, C) := \frac{c}{1-\beta^0(c)}, v_1(D, C) := \frac{d}{1-\beta^0(d)}$ and $v_1(C, D) := \frac{b}{1-\beta^0(b)}$. The payoffs for player 2 are defined analogously.

Lemma 6.9. *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 **??**, 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 6.5, $\tilde{a} := \big((D, C), ..., (D, C), (C, C), (C, C), ...\big)$ is efficient given η. However, $v_1(\tilde{a}, \lambda) > v_2(\tilde{a}, \lambda)$, which contradicts $\eta_1 < \eta_2$. Thus, $a^T = (C, D)$. By Lemma 6.8, the play path $\mathbf{a}' = \big((D, C), (C, D), (D, C), (C, D), ...\big) \in P(\lambda, \eta^{T-1})$ 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, $a^0 \neq (D, C)$. □

Let $\mathbf{a}(0) := \mathbf{a}^C$. Let $\mathbf{a}(T)$ be the path such that $a^t = (C, D)$ for all $0 \leq t \leq T - 1$ and ${}_T\mathbf{a} = \mathbf{a}^C$, where $T \geq 1$. Let $s(\mathbf{a}, \lambda)$ be the sum of payoffs, i.e., $s(\mathbf{a}, \lambda) = v_1(\mathbf{a}, \lambda) + v_2(\mathbf{a}, \lambda)$. The following lemma says that under $\eta = (1, 1)$, the path $\mathbf{a}(1)$ is never efficient.

Lemma 6.10. *For any* $\lambda \in (0, 1]$, $s(\mathbf{a}(1), \lambda) < \max\{s(\mathbf{a}^C, \lambda), s(\mathbf{a}^A, \lambda)\}$.

Proof. First compute the sum of payoffs:

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

By direct comparison, $s(\mathbf{a}^C, \lambda) > s(\mathbf{a}(1), \lambda)$ if and only if

$$\frac{c}{1-\beta^0(c)} > \frac{b+d}{1-\beta^0(b)+1-\beta^0(d)}. \tag{18}$$

Recall that $\beta = 1 - \lambda(1-\beta^0)$. Note that $s(\mathbf{a}^B, \lambda)$ decreases as λ decreases since $\frac{ds(\mathbf{a}^B,\lambda)}{d\lambda} > 0$. Since $s(\mathbf{a}^B, \lambda)$ decreases as λ decreases,

$$s(\mathbf{a}^B, \lambda) > \lim_{\lambda \to 0} s(\mathbf{a}^B, \lambda) = \frac{2(b+d)}{1-\beta^0(d)+1-\beta^0(b)}, \ \forall \lambda \in (0,1]. \tag{19}$$

When (18) holds with equality, $s(\mathbf{a}^A, \lambda) > s(\mathbf{a}(1), \lambda) = s(\mathbf{a}^C, \lambda)$, where the last equality follows (19). When the inequality in (18) is reversed, it is straightforward to verify that $s(\mathbf{a}^A, \lambda) > s(\mathbf{a}(1), \lambda)$. □

Proof of Theorem 5.1. Fix $\lambda \in (0,1]$. First we show that when $\eta = (1,1)$, any efficient path is one of dynamic cooperation. Take any $\mathbf{a} \in P(\lambda, \eta)$. By Lemma 6.4, (D, D) will never be chosen on any efficient path because it is strictly dominated by (C, C). Hence $a^0 \in \{(C,C), (D,C), (C,D)\}$. 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 $a^0 = (C, D)$. We only consider the following two cases: $a^0 = (C, C)$ and $a^0 = (C, D)$.

If $a^0 = (C, C)$, by Lemma 6.5, the path $\mathbf{a}^C$ is efficient given η. Consider the case in which $a^0 = (C, D)$. Since $\beta(b) > \beta(d)$, $\frac{\eta_1^1}{\eta_2^1} = \frac{\eta_1 \beta(b)}{\eta_2 \beta(d)} > 1$. By Lemma 6.9, $a^1 \in \{(C,C), (D,C)\}$. If $a^1 = (D, C)$, by Lemma 6.8, the path $\mathbf{a}^A$ is efficient. If $a^1 = (C, C)$, by Lemma 6.5, the path $\mathbf{a}(1)$ is efficient. Moreover, any efficient path $\hat{a}$ with $\hat{a}^0 = (C, D)$ and $\hat{a}^1 = (C, C)$ yields the same sum of payoffs as $\mathbf{a}(1)$. However, by Lemma 6.10, the path $\mathbf{a}(1)$ is always dominated by either $\mathbf{a}^C$ or $\mathbf{a}^A$, and hence $\hat{a}$ is also dominated. Therefore, $a^1 = (D, C)$.

We have found two efficient paths: $\mathbf{a}^C$ and $\mathbf{a}^A$. Intratemporal cooperation $\mathbf{a}^C$ is efficient if and only if inequality (**??**) holds. Otherwise, intertemporal cooperation is efficient.

Generically, if (C,C) is chosen on some efficient play path, then (C,C) should be chosen in every period, because no other path can yield the same weighted sum of payoffs. Conversely, if on some efficient play path there exists $a^t \neq (C,C)$, then (C,C) will not be chosen in any period. [2] Moreover, whenever $\eta_1^t = \eta_2^t$, we can choose $a^t \in \{(D,C),(C,D)\}$, as long as the next period $a^{t+1} \in \{(D,C),(C,D)\} \setminus a^t$. For example, the path $\big((C,D),(D,C),(D,C),(C,D),(C,D),(D,C),...\big)$ is efficient. Therefore, we have found all the efficient paths.

Next we characterize the efficient path for any given direction.

"Only if" part: Take $\eta \in \mathbb{R}^2_{++}$ such that $\eta_1 \neq \eta_2$. 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 6.9, we have $v_1(\mathbf{a},\lambda) \leq v_2(\mathbf{a},\lambda)$. By Lemma 6.4, (D,D) will never be chosen on any efficient path. By Lemma 6.9, $a^0 \neq (D,C)$. Hence, $a^0 \in \{(C,C),(C,D)\}$.

Suppose $a^0 = (C,C)$. By Lemma 6.5, the constant path $\big((C,C),(C,C),...\big)$ is efficient. It implies that (**??**) doesn't hold. If $\mathbf{a} = \big((C,C),(C,C),...\big)$, the efficient play path is $\mathbf{a}^C$ from time 0. Next we will show that if $a^0 = (C,C)$, it is impossible that $\mathbf{a} \neq \big((C,C),(C,C),...\big)$. Suppose $\mathbf{a} \neq \big((C,C),(C,C),...\big)$ 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 6.9, since $\frac{\eta_1}{\eta_2} < 1$, $a^T \neq (D,C)$. Thus, $a^T = (C,D)$. Since $\beta_1(b) > \beta_2(d)$, we have $\frac{\eta_1^{T+1}}{\eta_2^{T+1}} = \frac{\eta_1\beta_1(b)}{\eta_2\beta_2(d)} > \frac{\eta_1}{\eta_2}$. By symmetry, the Pareto frontier corresponding to the direction (η_2,η_1) is a linear segment connecting $\big(v_1(C,C),v_2(C,C)\big)$ and $\big(v_2({}_T\mathbf{a},\lambda),v_1({}_T\mathbf{a},\lambda)\big)$. 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 ${}_{T+1}\mathbf{a} = \big((C,C),(C,C),...\big)$. By Lemma 6.5, $v({}_T\mathbf{a},\lambda)$ and $v(C,C)$ are on the same linear segment of Pareto frontier, and η has the same direction as

$$\begin{aligned}&\big(v_1(C,C) - v_1({}_T\mathbf{a},\lambda), v_2({}_T\mathbf{a},\lambda) - v_2(C,C)\big)\\ &= \lambda\Bigg(\big(1-\beta^0(b)\big)\frac{c}{1-\beta^0(c)} - b, d - \big(1-\beta^0(d)\big)\frac{c}{1-\beta^0(c)}\Bigg).\end{aligned} \tag{20}$$

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

$$\eta_1(\lambda b + \beta(b)\frac{c}{1-\beta^0(c)}) + \eta_2(\lambda d + \beta(d)\frac{c}{1-\beta^0(c)}) = \eta_1\frac{c}{1-\beta^0(c)} + \eta_2\frac{c}{1-\beta^0(c)}. \tag{21}$$

[2]The special case is when $s(\mathbf{a}^A,\lambda) = s(\mathbf{a}^B,\lambda)$, any mixture of $\mathbf{a}^C$ and $\mathbf{a}^A$ can also be efficient.

Equation (20) and (21) hold if and only if

$$\big(1-\beta^0(b)\big)\frac{c}{1-\beta^0(c)} - b = d - \big(1-\beta^0(d)\big)\frac{c}{1-\beta^0(c)},$$

which implies $\eta_1 = \eta_2$. Thus, 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 6.8, the alternating path $\big((C, D), (D, C), (C, D), (D, C), ...\big)$ is efficient given η. By Lemma 6.5, $v(\mathbf{a}^A, \lambda)$ and $v(C, C)$ are on the same linear segment of Pareto frontier, and η has the same direction as

$$\begin{aligned}&\big(v_1(C, C) - v_1(\mathbf{a}^A, \lambda), v_2(\mathbf{a}^A, \lambda) - v_2(C, C)\big)\\ &= \Big(\frac{c}{1-\beta^0(c)} - \frac{\lambda(b+\beta(b)d)}{1-\beta(d)\beta(b)}, \frac{\lambda(d+\beta(d)b)}{1-\beta(d)\beta(b)} - \frac{c}{1-\beta^0(c)}\Big).\end{aligned} \tag{22}$$

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

$$\eta_1 \frac{\lambda(b+\beta(b)d)}{1-\beta(d)\beta(b)} + \eta_2 \frac{\lambda(d+\beta(d)b)}{1-\beta(d)\beta(b)} = \eta_1 \frac{c}{1-\beta^0(c)} + \eta_2 \frac{c}{1-\beta^0(c)}. \tag{23}$$

Both (22) and (23) hold if and only if

$$\frac{\lambda(b+\beta(b)d)}{1-\beta(d)\beta(b)} + \frac{\lambda(d+\beta(d)b)}{1-\beta(d)\beta(b)} = \frac{2c}{1-\beta^0(c)}. \tag{24}$$

However, if (24) holds, then we have $\eta_1 = \eta_2$, which contradicts our assumption that $\eta_1 < \eta_2$. Therefore, if $a^0 = (C, C)$, the unique efficient path is $\mathbf{a} = \mathbf{a}^C$. It also implies a stronger result, that if for some efficient play path $\mathbf{a}$ such that $a^T = (C, C)$ for some T, then the unique efficient continuation play path is ${}_T\mathbf{a} = \mathbf{a}^C$.

Suppose $a^0 = (C, D)$. By Lemma 6.7, the constant path $\big((C, D), (C, D), ...\big)$ 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 ${}_T\mathbf{a} = \mathbf{a}^C$. If $a^T = (D, C)$, the play path ${}_{T-1}\mathbf{a} = \mathbf{a}^A$ is efficient. This is the unique play path, because by Lemma 6.9, $a^{T+1} \neq (D, C)$, and by the argument above, if $a^{T+1} = (C, C)$, the constant path $\mathbf{a}^C$ will be the unique play path, which contradicts ${}_{T-1}\mathbf{a} = \mathbf{a}^A$ 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 $\mathbf{a}^C$ or $\mathbf{a}^A$. If equation (**??**) holds, geometrically it means the pair of payoffs from the path $\mathbf{a}^A$ is inside the Pareto frontier. Therefore, the continuation path can only be the constant path $\mathbf{a}^C$. If equation (**??**) doesn't hold, the pair of payoffs from constant path $\mathbf{a}^C$ is inside the Pareto frontier. Given any η, $\mathbf{a}^C$ cannot be efficient. So in this case, the continuation path can only be the alternating path $\mathbf{a}^A$.

"If" part: For this direction, we prove the following statement:

If inequality (**??**) holds, for any $T \geq 0$, there exists $\eta(T) \in \mathbb{R}^2_+$ such that $\mathbf{a}(T) \in P(\lambda, \eta(T))$.

Proof. Let $\eta(0) = (1, 1)$. We know that if $s(\mathbf{a}^C, \lambda) > s(\mathbf{a}^A, \lambda)$ holds, $\mathbf{a}(0) \in P(\lambda, \eta(0))$. For any $T \geq 1$, define $\eta(T)$ as follows

$$\begin{aligned}\eta(T) &= \Big(v_2(\mathbf{a}(T)) - v_2(\mathbf{a}(T-1)), v_1(\mathbf{a}(T-1)) - v_1(\mathbf{a}(T)\Big) \\ &= \Big(\lambda\beta^T(d)\big(d + \big(\beta(d) - 1\big)\frac{c}{1-\beta(c)}\big), \lambda\beta^T(b)\big(\big(1-\beta(b)\big)\frac{c}{1-\beta(c)} - b\big)\Big).\end{aligned}$$

The proof is shown by induction on T. We want to show that given any $T \geq 0$, if $\mathbf{a}(T) \in P(\lambda, \eta(T))$, then $\mathbf{a}(T+1) \in P(\lambda, \eta(T+1))$. From the proof above, we know that all the pure efficient paths should be in the form of $\mathbf{a}(T')$, for some $T' \geq 0$. If $\mathbf{a}(T+1) \in P(\lambda, \eta(T+1))$, it means

$$\eta_1(T+1)v_1(\mathbf{a}(T+1)) + \eta_2(T+1)v_2(\mathbf{a}(T+1)) \geq \eta_1(T+1)v_1(\mathbf{a}(T')) + \eta_2(T+1)v_2(\mathbf{a}(T')), \tag{25}$$

for all T'. When $T' > T+1$, inequality (25) is equivalent to

$$\beta(d) + ... + \beta^{T'-T-1}(d) < \beta(b) + ... + \beta^{T'-T-1}(b).$$

By IMI, $\beta(d) < \beta(b)$. Thus, the path $\mathbf{a}(T') \notin P(\lambda, \eta(T+1))$ for all $T' > T+1$. Note that by the construction of $\eta(T+1)$, we have

$$\eta_1(T+1)v_1(\mathbf{a}(T)) + \eta_2(T+1)v_2(\mathbf{a}(T)) = \eta_1(T+1)v_1(\mathbf{a}(T+1)) + \eta_2(T+1)v_2(\mathbf{a}(T+1)).$$

By assumption, $\mathbf{a}(T) \in P(\lambda, \eta(T))$. Since $\frac{\eta_1(T+1)}{\eta_2(T+1)} < \frac{\eta_1(T)}{\eta_2(T)}$, it implies (25) holds for all $T' < T$. Thus, the path $\mathbf{a}(T+1)$ is efficient. Since $\mathbf{a}(0) \in P(\lambda, \eta(0))$, by induction, $\mathbf{a}(T) \in P(\lambda, \eta(T))$ for all $T \geq 0$. □

Analogously, if inequality (**??**) holds, for any path that plays (D, C) in the first $T \geq 0$ periods and continues with $\mathbf{a}^C$, there exist $\eta \in \mathbb{R}^2_+$ such that this path is efficient. If inequality (**??**) doesn't hold, for any path that plays (C, D) (or (D, C)) in the first $T \geq 0$ periods and continues with $\mathbf{a}^A$, there exist $\eta \in \mathbb{R}^2_+$ such that this path is efficient. □

Proof of Corollary 5.1. Take any $\mathbf{a} \in P(\lambda, \eta) \cap IR^\varepsilon(\lambda)$. We need to show that $v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$, for all $t \in \mathcal{T}$ and $i = 1, 2$. From Theorem 5.1, we can see that there are only three possible efficient play paths: first, the constant path $\mathbf{a}^C$; second, start with (C, D) (or (D, C)) and after some finite time of periods continue with the constant path $\mathbf{a}^C$; last, start with (C, D) (or (D, C)) and after some finite time of periods continue with the alternating path $\mathbf{a}^A$. Geometrically it says if the initial pair of payoffs is not on the 45° line, it will move towards to the 45° line, and stay there afterwards. Given the assumption that the initial payoffs are strictly individually rational, the continuation payoffs at any time period will be strictly individually rational.

From the proof of Theorem 5.1, we know that $a^0 \in \{(C, C), (C, D)\}$. If $a^0 = (C, C)$, the unique efficient play path is $\mathbf{a}^C$. $v_i(\mathbf{a}, \lambda) \geq \varepsilon$ implies $v_i({}_t\mathbf{a}, \lambda) \geq \varepsilon$ for any t. Suppose $a^0 = (C, D)$. Let T be the first period such that $a^t \neq (C, D)$. For 0 to T, player 1's continuation payoff is increasing while player 2's continuation payoff is decreasing. If $a^T = (C, C)$ and as a result ${}_T\mathbf{a} = \mathbf{a}^C$, we have $v_1({}_t\mathbf{a}, \lambda) > v_1(\mathbf{a}, \lambda) \geq \varepsilon$ and $v_2({}_t\mathbf{a}, \lambda) \geq v_1({}_t\mathbf{a}, \lambda) \geq \varepsilon$ for all t. If $a^T = (D, C)$, the unique efficient play path starting from $T - 1$ is ${}_{T-1}\mathbf{a} = \mathbf{a}^A$. Since $v_1({}_t\mathbf{a}, \lambda) > v_1(\mathbf{a}, \lambda) \geq \varepsilon$, we have $v_2({}_t\mathbf{a}, \lambda) \geq v_2({}_T\mathbf{a}, \lambda) = v_1({}_{T-1}\mathbf{a}, \lambda) > \varepsilon$ for all t. Therefore, $\mathbf{a} \in SIR^\varepsilon$. □

Proof of Corollary 5.2. Take $\varepsilon > 0$. By Theorem 4.1, for any $\varepsilon > 0$, there exists $\overline{\lambda}$ such that for any $0 < \lambda < \overline{\lambda}$, any path $\mathbf{a} \in SIR^\varepsilon$ can be supported by an equilibrium. Take any $0 < \lambda < \overline{\lambda}$, $\eta \in \mathbb{R}^2_+$ and $\mathbf{a} \in P(\lambda, \eta)$. By Corollary 5.1, to show some path $\mathbf{a} \in SIR^\varepsilon(\lambda)$, it is sufficient to show $\mathbf{a} \in IR^\varepsilon(\lambda)$. Without loss of generality, assume inequality (**??**) holds and $a^t = (C, D)$ for $0 \leq t \leq T - 1$ and ${}_T\mathbf{a} = \mathbf{a}^C$. Then $\mathbf{a} \in IR^\varepsilon(\lambda)$ if and only if

$$v_1(\mathbf{a}, \lambda) = \frac{b}{1 - \beta^0(b)} + \beta^T(b)\left(\frac{c}{1 - \beta^0(c)} - \frac{b}{1 - \beta^0(b)}\right) \geq \varepsilon,$$

which is equivalent to

$$T \leq \frac{\ln\left(\varepsilon - \frac{b}{1-\beta^0(b)}\right) - \ln\left(\frac{c}{1-\beta^0(c)} - \frac{b}{1-\beta^0(b)}\right)}{\ln \beta(b)} := T(\varepsilon, \lambda).$$

Let T^* be the largest integer smaller than $T(\varepsilon,\lambda)$. By the characterization of the Pareto frontier, we have $\mathbf{a}(T^*+1), \mathbf{a}(T^*) \in P(\lambda, \eta(\varepsilon,\lambda))$, where

$$\eta(\varepsilon,\lambda) = \Big(v_1\big(\mathbf{a}(T^*),\lambda\big) - v_1\big(\mathbf{a}(T^*+1),\lambda\big), v_2\big(\mathbf{a}(T^*+1),\lambda\big) - v_2\big(\mathbf{a}(T^*),\lambda\big)\Big) \tag{26}$$

$$= \Big(\beta^{T^*}(b)\big(1-\beta^0(b)\big)\big(\frac{c}{1-\beta^0(c)} - \frac{b}{1-\beta^0(b)}\big), \beta^{T^*}(d)\big(1-\beta^0(d)\big)\big(\frac{d}{1-\beta^0(d)} - \frac{c}{1-\beta^0(c)}\big)\Big). \tag{27}$$

For any $\frac{\eta(\varepsilon,\lambda)_1}{\eta(\varepsilon,\lambda)_2} < \frac{\eta_1}{\eta_2} \leq 1$ and $\mathbf{a}(T) \in P(\lambda,\eta)$, we have $T \leq T^*$ and $\mathbf{a}(T) \in IR^\varepsilon(\lambda)$. Similar argument applies to $1 \leq \frac{\eta_1}{\eta_2} < \frac{\eta(\varepsilon,\lambda)_2}{\eta(\varepsilon,\lambda)_1}$. Therefore, for any $\frac{\eta(\varepsilon,\lambda)_1}{\eta(\varepsilon,\lambda)_2} < \frac{\eta_1}{\eta_2} < \frac{\eta(\varepsilon,\lambda)_2}{\eta(\varepsilon,\lambda)_1}$, we have $\mathbf{a} \in IR^\varepsilon(\lambda)$.

Also $(\frac{\eta(\varepsilon,\lambda)_1}{\eta(\varepsilon,\lambda)_2}, \frac{\eta(\varepsilon,\lambda)_2}{\eta(\varepsilon,\lambda)_1})$ is the largest interval this result can hold, because given $\eta(\varepsilon,\lambda)$, $\mathbf{a}(T^*) \in P(\lambda, \eta(\varepsilon,\lambda))$ but $\mathbf{a}(T^*) \notin IR^\varepsilon(\lambda)$. So for any $\frac{\eta(\varepsilon,\lambda)_1}{\eta(\varepsilon,\lambda)_2} \geq \frac{\eta_1}{\eta_2}$, on any efficient path $\mathbf{a}(T)$, we have $T \geq T^*+1$, and player 1's payoff will be less than ε; similarly, for any $\frac{\eta_1}{\eta_2} \geq \frac{\eta(\varepsilon,\lambda)_2}{\eta(\varepsilon,\lambda)_1}$, player 2's payoff will be less than ε. Thus, we have found the largest interval in which given any direction, the efficient path can be supported in equilibrium. □

Lemma 6.11. *Given any $\lambda \in (0,1]$, $\eta \in \mathbb{R}^2_+$ and $\mathbf{a} \in P(\lambda,\eta)$, under DMI, if $v_i(\mathbf{a},\lambda) = v_j(\mathbf{a},\lambda)$, then $\mathbf{a} = (\bar{a}^e, \bar{a}^e, ...)$.*

Proof. Suppose $v_i(\mathbf{a},\lambda) = v_j(\mathbf{a},\lambda)$ and $\beta_i(a^0) \neq \beta_j(a^0)$. 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 6.3, $v_i({}_1\mathbf{a},\lambda) \leq v_i(\mathbf{a},\lambda)$. By equation (14), $v_i(\mathbf{a},\lambda)$ 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 $a \in A^E$. By Lemma 6.4, $a = \bar{a}^e$. □

Lemma 6.12. *Given any $\lambda \in (0,1]$, $\eta \in \mathbb{R}^2_+$ and $\mathbf{a} \in P(\lambda,\eta)$, under DMI, if $v_i(\mathbf{a},\lambda) < v_j(\mathbf{a},\lambda)$, then $\beta_i(a^0) \leq \beta_j(a^0)$.*

Proof. Suppose $v_i(\mathbf{a},\lambda) < v_j(\mathbf{a},\lambda)$ and $\beta_i(a^0) > \beta_j(a^0)$. By DMI, $v_i(a^0) > v_j(a^0)$. If $v_j(a^0) \geq v_j(\mathbf{a},\lambda)$, then $\mathbf{a}$ is strictly Pareto dominated by constant play of a^0. So $v_j(a^0) <$

$v_j(\mathbf{a},\lambda)$. Also because $\beta_i(a^0) > \beta_j(a^0)$, player j's relative weight decreases at 1. By Lemma 6.3, $v_j({}_1\mathbf{a},\lambda) \leq v_j(\mathbf{a},\lambda)$, but this contradicts equation (14), because $v_j(\mathbf{a},\lambda)$ is a convex combination of $v_j(a^0)$ and $v_j({}_1\mathbf{a},\lambda)$. Therefore, $\beta_i(a^0) \leq \beta_j(a^0)$. □

Proof of Theorem 5.2. Take $\lambda \in (0,1]$ and $\eta \in \mathbb{R}^2_+$. Let $\mathbf{a} \in P(\lambda,\eta)$.

If $\eta_i = 0$ and $\eta_j > 0$, 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, $v_i(\mathbf{a},\lambda) = v_j(\mathbf{a},\lambda)$. From Lemma 6.11, $\mathbf{a} = (\overline{a}^e, \overline{a}^e, ...)$, that is, $a^t \in A^E$ for all $t \in \mathcal{T}$.

Second, consider the case $v_i(\mathbf{a},\lambda) < v_j(\mathbf{a},\lambda)$. By Lemma 6.12, $\beta_i(a^0) \leq \beta_j(a^0)$. Suppose $\beta_i(a^0) = \beta_j(a^0)$. 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 6.5 repeatedly and we have for all $0 < t \leq T'$, $\big(v_i({}_t\mathbf{a},\lambda), v_j({}_t\mathbf{a},\lambda)\big)$ 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 $0 < t \leq T'$. Apply Lemma 6.12 again, we know $\beta_i(a^{T'}) < \beta_j(a^{T'})$. Since $\frac{\eta_i^{T'+1}}{\eta_j^{T'+1}} = \frac{\eta_i\beta_i(a^{T'})}{\eta_j\beta_j(a^{T'})} < \frac{\eta_i}{\eta_j}$, we have $v_i({}_{T'+1}\mathbf{a},\lambda) \leq v_i(\mathbf{a},\lambda) < v_j(\mathbf{a},\lambda) \leq v_j({}_{T'+1}\mathbf{a},\lambda)$. Combining Lemma 6.6 and 6.12 yields $\beta_i(a^{T'+1}) < \beta_j(a^{T'+1})$. Repeat this process, for any $t \geq T'$, we have $\beta_i(a^t) < \beta_j(a^t)$.

Suppose $v_i(\mathbf{a},\lambda) < v_j(\mathbf{a},\lambda)$ and $\beta_i(a^0) < \beta_j(a^0)$. We argue that for any $t > 0$, $\beta_i(a^t) < \beta_j(a^t)$. Suppose not. Let T' be the first period such that $\beta_i(a^t) \geq \beta_j(a^t)$. Since $\frac{\eta_i^{T'}}{\eta_j^{T'}} < \frac{\eta_i}{\eta_j}$, any path $\hat{\mathbf{a}} \in P(\lambda, \eta^{T'})$ should satisfy $v_i(\hat{\mathbf{a}},\lambda) \leq v_i(\mathbf{a},\lambda) < v_j(\mathbf{a},\lambda) \leq v_j(\hat{\mathbf{a}},\lambda)$. Lemma 6.12 implies that $\beta_i(a^{T'}) = \beta_j(a^{T'})$ and $v_i(a^{T'}) = v_j(a^{T'})$. From Lemma 6.5, the path $\mathbf{a}' := (a^{T'}, a^{T'}, ...) \in P(\lambda, \eta^{T'})$. However, $v_i(\mathbf{a}',\lambda) = v_j(\mathbf{a}',\lambda)$, which is a contradiction. Hence, fro any $t > 0$, $\beta_i(a^t) < \beta_j(a^t)$. Hence, there exists some T large enough such that player j's relative weight is almost infinity. As a result, for any $t \geq T$, $a^t \in \overline{A}^j$. □

Proof of Corollary 5.3. First we show that in prisoners' dilemma, there are only three possible efficient play paths: $\big((C,D),(C,D),...\big)$, $\big((C,C),(C,C),...\big)$ and $\big((D,C),(D,C),...\big)$. Given λ, $\eta \in \mathbf{R}^2_+$ and $\mathbf{a} \in P(\lambda,\eta)$, if $v_1(\mathbf{a},\lambda) = v_2(\mathbf{a},\lambda)$, by Lemma 6.11, we have $\mathbf{a} = \big((C,C),(C,C),...\big)$.

If $v_1(\mathbf{a},\lambda) < v_2(\mathbf{a},\lambda)$, by Lemma 6.12, we have $\beta_1(a^0) \leq \beta_2(a^0)$. Suppose $\beta_1(a^0) = \beta_2(a^0)$, i.e., $a^0 = (C,C)$. Since $v_1(\mathbf{a},\lambda) < v_2(\mathbf{a},\lambda)$, let T be the first period such that $\beta_1(a^t) \neq \beta_2(a^t)$. Since $v_1(\mathbf{a},\lambda) < v_2(\mathbf{a},\lambda)$, by efficiency, we have $v_1(\mathbf{a},\lambda) < v_i(C,C) < v_2(\mathbf{a},\lambda)$, where $v_i(C,C) = \frac{c}{1-\beta^0(c)}$ and $i = 1,2$. By equation (14), $v_1({}_t\mathbf{a},\lambda) < v_1(\mathbf{a},\lambda) < v_1(C,C) < v_2(\mathbf{a},\lambda) < v_1({}_t\mathbf{a},\lambda)$, for all $0 < t \leq T$. By Lemma 6.12, we have $a^T = (C,D)$. From Lemma 6.6, $a^t = (C,D)$ for all $t \geq T$. Hence $v_1({}_T\mathbf{a},\lambda) = v_1({}_{T+1}\mathbf{a},\lambda)$. By equation (14), $v_1({}_T\mathbf{a},\lambda)$ is a convex combination of $v_1({}_{T+1}\mathbf{a},\lambda)$ and $v_1(C,C)$, which contradicts $v_1({}_T\mathbf{a},\lambda) = v_1({}_{T+1}\mathbf{a},\lambda) < v_1(C,C)$. Therefore, if $v_1(\mathbf{a},\lambda) < v_2(\mathbf{a},\lambda)$, we have $a^0 = (C,D)$. Since $\beta_1(a^0) < \beta_2(a^0)$, player 1's relative weight decreases. Hence, $v_1({}_1\mathbf{a},\lambda) \leq v_1(\mathbf{a},\lambda) < v_2(\mathbf{a},\lambda) \leq v_1({}_1\mathbf{a},\lambda)$. As a result, $a^1 = (C,D)$. Apply this argument repeatedly, and we get $\mathbf{a} = \big((C,D),(C,D),...\big)$.

The case in which $v_1(\mathbf{a},\lambda) > v_2(\mathbf{a},\lambda)$ is symmetric to the case above, and the efficient play path is $\big((D,C),(D,C),...\big)$. Therefore, any efficient path is one of these three paths or a mixture of them.

Since $b < 0$, the path $\big((C,D),(C,D),...\big)$ and $\big((D,C),(D,C),...\big)$ cannot arise in any equilibrium. Thus, the only possible efficient path that can arise in an equilibrium is the constant path $\big((C,C),(C,C),...\big)$. If (8) doesn't hold, it means this path is dominated by a mixture of $\big((C,D),(C,D),...\big)$ and $\big((D,C),(D,C),...\big)$ with equal probability. If (8) holds, then $\big((C,C),(C,C),...\big)$ is the unique efficient path that can arise in an equilibrium. □

6.2 Prinsoners' Dilemma Numerical Example

	C	D
C	$2,2$	$-1,3$
D	$3,-1$	$0,0$

Figure 2: The prisoners' dilemma numerical example

Let $\beta^0(d) = 0.7, \beta^0(b) = 0.8, \beta^0(c) = 0.75$. The sum of payoffs from the constant play of (C,C) is $2\frac{c}{1-\beta^0(c)} = 16$. The sum of payoffs from alternating between (C,D) and (D,C) is at most $\frac{d+\beta^0(d)b}{1-\beta^0(d)\beta^0(b)} + \frac{b+\beta^0(b)d}{1-\beta^0(d)\beta^0(b)} = \frac{185}{22} < 16$, which is smaller than the sum of payoffs from the constant path $\big((C,C),(C,C),...\big)$. If $\eta_1 < \eta_2$, by applying Theorem 5.1, the efficient path will be $\big((C,D),...,(C,D),(C,C),(C,C),...\big)$, that is, playing (C,D) in the first T periods, and then playing (C,C) thereafter. For each initial direction and each $\lambda \in (0,1]$, there exists an optimal T.

Next we show that there exists a non-trivial set of directions in which the conditions in Corollary 5.2 hold, that is, there exist $\varepsilon > 0$ and $\eta \in \mathbb{R}^2_+$ such that for any $\mathbf{a} \in P(\lambda, \eta)$, $v_i(\mathbf{a}, \lambda) \geq \varepsilon, i = 1, 2$, for all $\lambda \in (0, 1]$.

Proof. Take $\eta \in \mathbb{R}^2_{++}$ and $\eta_1 < \eta_2$. Fix $\lambda \in (0, 1]$. By Theorem 5.1, the efficient play path is $\big((C, D), ..., (C, D), (C, C), (C, C), ...\big)$, that is, playing (C, D) in the first T periods, and then playing (C, C) thereafter. For each initial direction and each $\lambda \in (0, 1]$, there exists an optimal T. The maximization problem becomes

$$\max \eta_1 v_1 + \eta_2 v_2,$$

where $v_1 = \frac{b\left(1-\beta^T(b)\right)}{1-\beta^0(b)} + \beta^T(b)\frac{c}{1-\beta^0(c)}$ and $v_2 = \frac{d\left(1-\beta^T(d)\right)}{1-\beta^0(d)} + \beta^T(d)\frac{c}{1-\beta^0(c)}$. Take first order condition with respect to T, we get

$$T^* = \frac{\ln \frac{\eta_2\left(\frac{d}{1-\beta^0(d)} - \frac{c}{1-\beta^0(c)}\right) \ln \beta(d)}{\eta_1\left(\frac{c}{1-\beta^0(c)} - \frac{b}{1-\beta^0(b)}\right) \ln \beta(b)}}{\ln \frac{\beta(b)}{\beta(d)}}. \tag{28}$$

We want to find the directions in which both players' payoffs are greater than zero. Because $\eta_1 < \eta_2$, player 1's payoff is smaller than player 2's payoff. It is sufficient to check whether player 1's payoff is bounded above 0 for all $\lambda \in (0, 1]$. Player 1's payoff $v_1 = \frac{b\left(1-\beta^T(b)\right)}{1-\beta^0(b)} + \beta^T(b)\frac{c}{1-\beta^0(c)} > 0$ if and only if

$$T < \frac{\ln \frac{-\frac{b}{1-\beta^0(b)}}{\frac{c}{1-\beta^0(c)} - \frac{b}{1-\beta^0(b)}}}{\ln \beta(b)}. \tag{29}$$

We need to find the directions such that T^* defined in (28) satisfies (29), for all $\lambda \in (0, 1]$. It means given these directions, player 1's efficient payoff is individually rational for all $\lambda \in (0, 1]$. Plug in the numbers given above, direct verification shows that when $\frac{\eta_2}{\eta_1} \geq 7$, there exist some λ such that T^* violates (29). When $1 < \frac{\eta_2}{\eta_1} \leq \frac{13}{2}$, T^* satisfies (29) for all $\lambda \in (0, 1]$. By symmetry, if $\frac{2}{13} \leq \frac{\eta_i}{\eta_j} \leq \frac{13}{2}$, there exists $\varepsilon > 0$ such that for any $\mathbf{a} \in P(\lambda, \eta)$, $v_i(\mathbf{a}, \lambda) \geq \varepsilon, i = 1, 2$, for all $\lambda \in (0, 1]$. □

Note that in the proof we give an explicit formula to compute T^*, which generally is not an integer. However, in any efficient play path, the optimal number of periods in which (C, D) is played should be an integer. Since the weighted sume of players' payoffs increases as $T < T^*$ and decreases as $T > T^*$, the optimal number of periods in which (C, D) is played should be the closest integer to T^*. For example, when $\frac{\eta_2}{\eta_1} = \frac{13}{2}$ and take $\lambda = 0.1$, $T^* = 40.03$. Thus, the optimal number of periods in which (C, D) is played is 40, that is, the efficient path is playing (C, D) for 40 periods, and playing (C, C) thereafter.

Next we show that there exists η and λ, λ' such that for $\mathbf{a} \in P(\lambda, \eta)$ and $\mathbf{a}' \in P(\lambda', \eta)$, we have $v_1(\mathbf{a}, \lambda) > 0$ while $v_1(\mathbf{a}', \lambda') < 0$. Take $(\eta_1, \eta_2) = (1, 7)$, $\lambda = 0.2$ and $\lambda' = 0.02$. From equation (28), when $\lambda = 0.2$, the optimal number of periods in which (C, D) is played is 23, and direct verification shows that $v_1(\mathbf{a}, \lambda) > 0$. When $\lambda = 0.02$, the optimal number of periods in which (C, D) is played is 239, and direct verification shows that $v_1(\mathbf{a}', \lambda') < 0$. Thus, in the direction $(\eta_1, \eta_2) = (1, 7)$, the condition in Corollary 5.2 is not satisfied.

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

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

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

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

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

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

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

[11] D. Fudenberg and J. Tirole. Game theory, 1991. *Cambridge, Massachusetts*, 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.