Characterizing the Limit Set of PPE Payoffs with Unequal Discounting

Takuo Sugaya*†
Princeton University

February 10, 2010

Abstract

We study the repeated games with imperfect public monitoring and unequal discounting. We characterize the limit set of perfect and public equilibrium payoffs as discount factors converge to 1 with the relative patience between players fixed. We show that the pairwise full rank condition is sufficient for the folk theorem.

*tsugaya@princeton.edu

†The author is thankful to Stephen Morris and Satoru Takahashi for their helpful guidance. I am also grateful to Dilip Abreu, Chiaki Hara, Yuichiro Kamada, Akihiko Matsui, Larry Samuelson, Tadashi Sekiguchi, and Katsutoshi Wakai for their valuable comments and to participants of 2008 Summer Workshop in Economic Theory. The usual disclaimer of course applies.

1 Introduction

In this paper, we characterize the equilibrium payoffs in repeated games with imperfect public monitoring and unequal discounting. In particular, we firstly recursively characterize the set of feasible and sequentially individually rational (henceforth SIR) payoffs and that of perfect and public equilibrium (henceforth PPE) payoffs for fixed discount factors à la Abreu, Pearce, and Stacchetti (1990) (henceforth APS). Secondly, we characterize the limit sets of SIR and PPE payoffs respectively as discount factors converge to 1 with the relative patience between players fixed. As a corollary, we show that the pairwise full rank condition introduced by Fudenberg, Levine, and Maskin (1994) (henceforth FLM) is sufficient for the folk theorem with unequal discounting, that is, the limit set of PPE payoffs coincides with that of SIR payoffs.

The mostly related papers are Lehrer and Pauzner (1999) (henceforth LP), Fudenberg and Levine (1994) (henceforth FL), and FLM. LP analyze the repeated games with perfect monitoring and unequal discounting. They define SIR payoffs and show the folk theorem with two players. This paper extends their results into two directions. Firstly, we obtain the characterization and the folk theorem with imperfect public monitoring. Secondly, our results hold for general n-player games. With unequal discounting, the intertemporal trade is important to achieve extreme points and, in two-player games, the direction of the intertemporal trade is monotone. For example, for the Pareto frontier, it is efficient to play actions that the impatient player prefers in early stages and actions that the patient player prefers in later stages. A novelty of this paper is to overcome the difficulty from the non-monotonicity of the intertemporal trade in n-player games.

On the other hand, FL show that the limit set of PPE payoffs can be characterized by the solution of a family of static liner programming problems (FL problem) in equal discounting when the dimensionality condition is satisfied. FLM show that if the pairwise full rank condition is satisfied, this characterized set coincides with the feasible and individually rational payoff set. This paper shows that the dimensionality condition of FL problem is sufficient to attain the limit characterization with unequal discounting and that the pairwise full rank condition is sufficient to attain the folk theorem with unequal discounting.

Other related papers are Harrington (1989), Haag and Lagunoff (2007), and Salonen and Vartiainen (2008). All of them consider perfect monitoring while we consider imperfect monitoring. Haag and Lagunoff (2007) consider stationary equilibria while we consider non-stationary equilibria taking the intertemporal trade into account. Salonen and Vartiainen (2008) construct an example such that without public randomization, the set of PPE payoffs in repeated games with unequal discount factors is not convex (or even not connected). In this paper, we show that this is not the case in the limit, that is, the limit set of PPE payoffs is convex without public randomization.

The rest of the paper is organized as follows. Section 2 defines the model. In Section 3, we characterize the set of SIR payoffs and the set of PPE payoffs in a recursive way. Section 4 is devoted to our main results; we give the limit characterization of PPE payoffs that is valid if the solution for FL problem satisfies the full dimensionality condition. In Section 5, we show that the pairwise full rank condition implies the folk theorem. Section 6 discusses possible extensions and concludes.

2 Model

2.1 The Stage Game

We consider a stage game with n players, $1, 2, ..., n$. In the stage game, players move simultaneously and player i chooses an action a_i from a set A_i. We restrict our attention to a finite game, that is, $|A_i| < \infty$ for all i. Let $a \in A \equiv \prod_{i=1}^{n} A_i$ be an action profile. An action profile induces a probability distribution over a possible public outcome $y \in Y$, where Y is a finite set. Each player i's realized payoff $r_i(a_i, y)$ depends only on his action a_i and the public outcome y. Let $\rho(y \mid a)$ be the probability of y given a. Player i's expected payoff from a is given by $g_i(a) \equiv \sum_{y \in Y} \rho(y \mid a) r_i(a_i, y)$. Define $g(A) \equiv \{g(a)\}_{a \in A}$ for a notational convenience.

Letting $\mathcal{A}_i \equiv \triangle(A_i)$ be the set of probability distributions over A_i. Then, a mixed action α_i for each player i is an element of $\mathcal{A}_i$. Let $\alpha_i(a_i)$ be the probability that α_i assigns to a_i. Given $\alpha \equiv (\alpha_1, \alpha_2, ..., \alpha_n) \in \mathcal{A} \equiv \prod_{i=1}^{n} \mathcal{A}_i$, define $\rho(y \mid \alpha) \equiv \sum_{a \in A} \rho(y \mid a) \alpha(a)$ and $g_i(\alpha) \equiv \sum_{y \in Y} \sum_{a \in A} \rho(y \mid a) \alpha(a) r_i(a_i, y)$ with $\alpha(a) \equiv \alpha_1(a_1) \alpha_2(a_2) \cdots \alpha_n(a_n)$. As usual, we define $a_{-i} \equiv (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n) \in A_{-i} \equiv \prod_{j \neq i} A_j$. The similar convention will be used if it causes no confusion.

Let ir_i be the individually rational payoff level for player i: $ir_i \equiv \min_{\alpha_{-i} \in \mathcal{A}_{-i}} \max_{a_i \in A_i} g_i(a_i, \alpha_{-i})$. IR_i denotes the set of individually rational payoffs, that is, $IR_i \equiv \{v \in \mathbb{R}^n : v_i \geq ir_i\}$, and $IR \equiv \bigcap_{i=1}^{n} IR_i$. Note that the set of feasible payoffs of the stage game is $\text{co}(g(A))$ and the set of feasible and individually rational payoffs of the stage game is $\mathrm{co}\,(g\,(A)) \cap IR$.

2.2 The Repeated Game

The stage game is played infinitely many times and in each period $t = 1, 2, \ldots$, each player observes the resulting public signal y_t. The total payoffs of the players are defined as follows. Given a sequence of probability distributions over the stage game payoff vectors and players' discount factors $\boldsymbol{\delta} \equiv (\delta_1, \ldots, \delta_n)$, player i's utility in the repeated game is the discounted sum of his expected stage game payoffs. In keeping with the literature, it is equivalent to consider his expected discounted average payoff, that is, letting $\{g_i^t\}_{t=1}^{\infty}$ be player i's sequence of expected stage game payoffs, his total payoff is given by

$$(1 - \delta_i) \sum_{t=1}^{\infty} \delta_i^{t-1} g_i^t.$$

2.2.1 Feasible and Sequentially Individually Rational Payoffs

The payoff is feasible and sequentially individually rational (henceforth SIR) if it is attainable by a sequence of correlated actions with the continuation payoff of each player in each period greater than her individually rational payoff. Let $\mu \in \triangle(A)$ be a generic element of correlated actions, $\rho(y \mid \mu) \equiv \sum_{a \in A} \rho(y \mid a)\mu(a)$ be the probability distribution over public outcomes, and $g_i(\mu) \equiv \sum_{y \in Y} \sum_{a \in A} \rho(y \mid a)\mu(a) r_i(a_i, y)$ be player i's expected payoff, respectively. The formal definition of SIR payoffs is given as follows.

Definition 1 *A payoff profile v is feasible and sequentially individually rational (SIR) if there exists $\{\mu^t\}_{t=0}^{\infty}$ with $\mu^t \in \triangle(A)$ for all t such that, for all i,*

$$v_i = (1 - \delta_i) \sum_{t=1}^{\infty} \delta_i^{t-1} g_i(\mu^t),$$

$$v_i^{\tau} \equiv (1 - \delta_i) \sum_{t=\tau}^{\infty} \delta_i^{t-\tau} g_i(\mu^t) \geq ir_i \text{ for all } \tau \geq 1.$$

Let $F(\boldsymbol{\delta})$ be the set of SIR payoff profiles. As LP point out, if the discount factors are unequal, $F(\boldsymbol{\delta})$ may be larger than the set of feasible and individually rational payoffs of the stage game.[1] In

[1] For more discussions about SIR, see LP.

general,

$$F^{\square} \equiv \prod_{i=1}^{n} \left[ir_i, \max_{a \in A} g_i(a) \right] \supset F(\boldsymbol{\delta}) \supset \mathrm{co}(g(A)) \cap IR.$$

2.2.2 Perfect and Public Equilibrium

We restrict our attention to perfect and public equilibrium (PPE) in this paper. Since a_i is player i's private information and y is a public outcome, the public history at the beginning of period t is $h^t = (\emptyset, y_1, \ldots, y_{t-1})$ and player i's private history is $h_i^t = (\emptyset, a_{1,i}, \ldots, a_{t-1,i})$. The set of histories for player i is $\mathcal{H}_i \equiv \bigcup_{t=0}^{\infty} (A_i \times Y)^t$ and the set of public histories is $\mathcal{H} \equiv \bigcup_{t=0}^{\infty} Y^t$. Player i's public strategy is a mapping from $\mathcal{H}$ to $\mathcal{A}_i$.

Each strategy profile generates a probability distribution over the sequence of stage game payoff vectors. Player i's objective in the repeated game is to maximize his expected discounted average payoff, that is, letting $\{g_i^t\}_{t=1}^{\infty}$ be player i's sequence of expected stage game payoffs, he maximizes $(1-\delta_i)\sum_{t=1}^{\infty} \delta_i^{t-1} g_i^t$.

We concentrate on perfect public equilibrium (henceforth PPE), where player i's strategy σ_i is a public strategy and the strategy profile σ forms a Nash equilibrium after any public history. Let $E(\boldsymbol{\delta})$ be the set of PPE payoffs.

3 Recursive Characterization

In this section, we recursively characterize the set of SIR payoff profiles, $F(\boldsymbol{\delta})$, and that of PPE payoff profiles, $E(\boldsymbol{\delta})$, as Abreu, Pearce, and Stacchetti (1990) (henceforth APS). In the following sections, based on this recursive characterization, we characterize the limit set of SIR payoffs and that of PPE payoff set as discount factors converge to 1 with the relative patience fixed.

3.1 The SIR Payoffs

Firstly, we give a recursive characterization of $F(\boldsymbol{\delta})$. As LP pointed out, since $F(\boldsymbol{\delta})$ depends on discount factors with unequal discounting, this characterization is nontrivial.

The following two notions are useful.

Definition 2 (SIR decomposability) *$v \in IR$ is sequentially individually rationally (SIR) decomposable on $W \subset \mathbb{R}^n$ if there exist $\mu \in \triangle(A)$ and $w \in W \cap IR$ such that $v_i = (1-\delta_i)\, g_i(\mu) + \delta_i w_i$*

for all i. Let $\mathbf{B}^F(W:\boldsymbol{\delta})$ be the set of all SIR decomposable payoffs on $W \subset \mathbb{R}^n$, that is,

$$\mathbf{B}^F(W:\boldsymbol{\delta}) = \left\{ v \in IR : \begin{array}{c} \exists \mu \in \triangle(A) \text{ and } w \in W \cap IR \text{ such that} \\ v_i = (1-\delta_i) g_i(\mu) + \delta_i w_i \text{ for all } i \end{array} \right\}.$$

Definition 3 (SIR self-generating) *A set of payoffs $W \subset \mathbb{R}^n$ is SIR self-generating if $W \subset \mathbf{B}^F(W:\boldsymbol{\delta})$.*

In words, v is SIR-decomposable on W if there exist a probability distribution μ and $w \in W \cap IR$ such that if players take the correlated action μ and the continuation payoffs are given by w, the total payoff is equal to v. Note that we allow players to take a correlated action and require continuation payoffs to be in IR, which guarantees that $F(\boldsymbol{\delta})$ is the largest SIR self-generating set in $F^{\square}$.

Proposition 1 *$F(\boldsymbol{\delta})$ is the largest SIR self-generating set included in $F^{\square}$ and $F(\boldsymbol{\delta})$ is convex and compact.*

The proof of this and all subsequent results are in the Appendix.

3.2 The PPE Payoffs

Secondly, we give a recursive characterization of $E(\boldsymbol{\delta})$. Since a PPE preserves the recursive structure with unequal discounting, APS can be extended directly.

Definition 4 (enforceability) *For $v \in \mathbb{R}^n$, $\alpha \in \mathcal{A}$, and $\{w(y)\}_{y\in Y}$, $\{w(y)\}_{y\in Y}$ enforces $\langle v, \alpha \rangle$ if*

$$\begin{gathered} v_i = (1-\delta_i) g_i(\alpha) + \delta_i E[w_i(y):\alpha] = (1-\delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[w_i(y):a_i,\alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ v_i = (1-\delta_i) g_i(\alpha) + \delta_i E[w_i(y):\alpha] \geq (1-\delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[w_i(y):a_i,\alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0. \end{gathered}$$

Definition 5 (decomposability) *$v \in \mathbb{R}^n$ is decomposable on $W \subset \mathbb{R}^n$ if there exist $\alpha \in \mathcal{A}$ and $\{w(y)\}_{y\in Y}$ with $w(y) \in W$ for all y such that $\{w(y)\}_{y\in Y}$ enforces $\langle v, \alpha \rangle$. Let $\mathbf{B}(W:\boldsymbol{\delta})$ be the set of all decomposable payoff profiles on $W \subset \mathbb{R}^n$.*

Definition 6 (self-generating) *A set of payoffs* $W \subset \mathbb{R}^n$ *is self-generating if* $W \subset \mathbf{B}(W : \boldsymbol{\delta})$.

Proposition 2 $E(\boldsymbol{\delta})$ *is the largest self-generating set included in* $F^{\square}$ *and* $E(\boldsymbol{\delta})$ *is compact.*

4 Limit Characterization

In this section, based on the recursive characterization of PPE payoffs, we characterize the limit of the set of PPE payoffs $E(\boldsymbol{\delta})$. With unequal discounting, with discount factors converging to 1, we should keep the relative patience fixed for all the players. In this section, we fix the relative patience in a certain way, that is, we consider the limit of $\delta_i = 1/(1+r_i\varepsilon)$ for all i as ε converges to 0. This is equivalent to keeping $\frac{1-\delta_i}{\delta_i}/\frac{1-\delta_n}{\delta_n} = r_i/r_n$ for all i, which means the ratio of the relative importance of instantaneous utilities against continuation payoffs is constant. We can extend our results to the more general limit where we only require $(1-\delta_i)/(1-\delta_n) \to r_i$ and $\delta_i \to 1$ for all i. See Section 6 for more details. We normalize $r_1 \geq \cdots \geq r_n = 1$ and, for notational convenience, we define R as a diagonal matrix whose ith entry is r_i.

For equal discounting, FL show that the limit set of PPE payoffs can be characterized by the solution of a family of static linear programming problems (henceforth FL problem):

$$\lim_{\varepsilon \to 0} E(\boldsymbol{\delta}) = \bigcap_{\lambda \in \Lambda} H(\lambda) \text{ for all } \boldsymbol{\delta} \text{ with } \delta_i = 1/(1+\varepsilon) \text{ for all } i,$$

where

$$\begin{aligned}
\Lambda &\equiv \{\lambda \in \mathbb{R}^n : \|\lambda\| = 1\}, \\
H(\lambda) &\equiv \{v \in \mathbb{R}^n : \lambda \cdot v \leq \sup_{\alpha \in \mathcal{A}} k(\alpha, \lambda)\}, \\
k(\alpha, \lambda) &\equiv \max_{v \in \mathbb{R}^n, \{w(y)\}_{y \in Y}} \lambda \cdot v \text{ subject to} \\
&\left\{\begin{array}{l}
v_i = (1-\delta) g_i(a_i, \alpha_{-i}) + \delta E[w_i(y) : a_i, \alpha_{-i}] \\
\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
v_i \geq (1-\delta) g_i(a_i, \alpha_{-i}) + \delta E[w_i(y) : a_i, \alpha_{-i}] \\
\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
0 \geq \lambda \cdot (w(y) - v) \text{ for all } y \in Y,
\end{array}\right.
\Leftrightarrow
\left\{\begin{array}{l}
v_i = g_i(a_i, \alpha_{-i}) + E[x_i(y) : a_i, \alpha_{-i}] \\
\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
v_i \geq g_i(a_i, \alpha_{-i}) + E[x_i(y) : a_i, \alpha_{-i}] \\
\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
0 \geq \lambda \cdot x(y) \text{ for all } y \in Y,
\end{array}\right.
\end{aligned} \quad (1)$$

with $x(y) \equiv \frac{\delta}{1-\delta}(w(y) - v)$.

A natural extension of FL problem is to write down (1) with unequal discounting, which gives us

$$k(\alpha, \lambda) \equiv \max_{v \in \mathbb{R}^n, \{w(y)\}_{y \in Y}} \lambda \cdot v \text{ subject to}$$
$$\begin{cases} v_i = (1 - \delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[w_i(y) : a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ v_i \geq (1 - \delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[w_i(y) : a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\ 0 \geq \lambda \cdot (w(y) - v) \text{ for all } y \in Y. \end{cases}$$

However, this always gives us $k(\alpha, \lambda) = \infty$ if λ is not parallel to $R\lambda$.

To clarify the problem, let us consider the two-player case with $\lambda = (1/\sqrt{2}, 1/\sqrt{2})$ and $r_1 > r_2 = 1$. Suppose we have a bounded solution $(v^*, \{w^*(y)\}_{y \in Y})$. Then, $w_1(y) = w_1^*(y) - K$ and $w_2(y) = w_2^*(y) + K$ for all y satisfy all the conditions. The effect on $\lambda \cdot v$ is

$$\frac{1}{\sqrt{2}}(\delta_2 - \delta_1) K > 0.$$

The key observation is that since player 1 is less patient than player 2, the total effect of subtracting K from the continuation payoff of player 1 and giving it to player 2 is strictly positive. More generally, if we can find w such that

$$\begin{aligned} 0 &\geq \lambda \cdot w, \\ 0 &< \sum_{i=1}^{n} \lambda_i \delta_i w_i, \end{aligned}$$

we can increase $k(\alpha, \lambda)$ by replacing $w(y)$ with $w(y) + w$ and the existence of such w, which is called the gain from the intertemporal trade by LP, depends on the global shape of the limit of $E(\boldsymbol{\delta})$.

Therefore, we provide the *recursive* characterization of the limit of $E(\boldsymbol{\delta})$. Given W, calculate the following

$$
\begin{aligned}
& k(\lambda : W : R) \\
= & \sup_{v\in\mathbb{R}^n,\alpha\in\mathcal{A},\{w(y)\}_{y\in Y}} \lambda\cdot v \text{ subject to} \\
& \left\{\begin{array}{l}
v_i = (1-\delta_i)\, g_i(a_i,\alpha_{-i}) + \delta_i E[w_i(y) : a_i,\alpha_{-i}] \\
\text{for all } i \text{ and } a_i\in A_i \text{ such that } \alpha_i(a_i) > 0, \\
v_i \geq (1-\delta_i)\, g_i(a_i,\alpha_{-i}) + \delta_i E[w_i(y) : a_i,\alpha_{-i}] \\
\text{for all } i \text{ and } a_i\in A_i \text{ such that } \alpha_i(a_i) = 0, \\
0 \geq \lambda\cdot(w(y) - v) \text{ for all } y\in Y, \\
v\in W.
\end{array}\right.
\Leftrightarrow
\left\{\begin{array}{l}
v_i = g_i(a_i,\alpha_{-i}) + E[x_i(y) : a_i,\alpha_{-i}] \\
\text{for all } i \text{ and } a_i\in A_i \text{ such that } \alpha_i(a_i) > 0, \\
v_i \geq g_i(a_i,\alpha_{-i}) + E[x_i(y) : a_i,\alpha_{-i}] \\
\text{for all } i \text{ and } a_i\in A_i \text{ such that } \alpha_i(a_i) = 0, \\
0 \geq R\lambda\cdot x(y) \text{ for all } y\in Y, \\
v\in W.
\end{array}\right.
\end{aligned}
$$

with $x_i(y) \equiv \frac{\delta_i}{1-\delta_i}(w_i(y) - v_i)$. The condition that $v\in W$ implies that the gain from the intertemporal trade is restricted by the shape of W. Since the first three constraints are equivalent to FL problem with direction $R\lambda$,

$$k(\lambda : W : R) = \max_{v\in W\cap H(R\lambda)} \lambda\cdot v.$$

Define

$$
\begin{aligned}
H(\lambda : W : R) &\equiv \{v : \lambda\cdot v \leq k(\lambda : W : R)\}, \\
\mathcal{B}(W : R) &\equiv \bigcap_{\lambda\in\Lambda} H(\lambda : W : R).
\end{aligned}
$$

Since $\mathcal{B}(\cdot : R)$ is weakly decreasing, $W\subset F^{\square} \Rightarrow \mathcal{B}(W : R)\subset F^{\square}$. In addition, $\mathcal{B}(W : R)$ is convex, compact, and monotone. Therefore, there exists a largest fixed point of $\mathcal{B}(\cdot : R)\subset F^{\square}$ and any fixed point is convex and compact. Let Q^R be the largest fixed point of $\mathcal{B}(\cdot : R)\subset F^{\square}$.

We will show that $Q^R = \lim_{\varepsilon\to 0} E(\boldsymbol{\delta})$ under the assumption that $\dim\bigcap_{\lambda\in\Lambda} H(R\lambda) = n$.

Assumption 1 $\dim\bigcap_{\lambda\in\Lambda} H(R\lambda) = n$.

Several comments are necessary on **Assumption 1**. Firstly, $\dim\bigcap_{\lambda\in\Lambda} H(R\lambda) = \dim\bigcap_{\lambda\in\Lambda} H(\lambda)$, that is, **Assumption 1** is equivalent to assuming the full dimensionality for FL problem.

Secondly, when **Assumption 1** is violated, the exact shape of the set of PPE payoffs is an open question even with perfect monitoring if discounting is unequal. In particular, we do not know whether the folk theorem holds or not. The basic intuition of the necessity of the full dimensionality for equal discounting is as follows[2]: if the full dimensionality condition is violated, it implies that more than one players share the same preference. To give incentives to punish a player, we must give "carrots" for the other players after the punishment phase. However, if the punished player share the same preference with one of the punishers, the punished player also gets carrots, which reduces the severeness of the punishment. As Chen (2007) points out, with unequal discounting, however, even if the static preferences are the same, the intertemporal preferences are different. Therefore, it might be possible to attain the folk theorem without full dimensionality.[3]

Thirdly, it is common to assume the full dimensionality for FL problem in the literature: with perfect monitoring, Fudenberg and Maskin (1986) firstly introduce the assumption. Abreu, Dutta, and Smith (1994) relax the assumption and Wen (1994) characterizes the equilibrium payoff set when the full dimensionality condition is violated. With public monitoring, FLM and FL assume the condition. Fudenberg, Levine, and Takahashi (2007) characterize the equilibrium payoff set when the full dimensionality condition is not satisfied.

It might be possible to attain the tight characterization for the case without **Assumption 1** à la Fudenberg, Levine, and Takahashi (2007) but it is beyond the scope of this paper. Hence, we leave this for the future research and proceed to state out main theorem:

Theorem 1 *1. For all* $\boldsymbol{\delta}$ *with* $\delta_i = 1/(1 + r_i\varepsilon)$ *for all* i, $E(\boldsymbol{\delta}) \subset Q^R$.

2. If ***Assumption 1*** *is satisfied, for all* $\boldsymbol{\delta}$ *with* $\delta_i = 1/(1 + r_i\varepsilon)$ *for all* i, $\lim_{\varepsilon\to 0} E(\boldsymbol{\delta}) = Q^R$.

We will give the sketch of the second argument.

With equal discounting, the proof proceeds as follows.

1. Take some point $o \in \text{int}Q^I$. For $t \in (0, 1)$, take the radial contraction of Q^R with respect to o, that is, $Q_1^R \equiv \{v : \exists v' \in Q^R \text{ such that } v = (1 - t)v' + to\}$. We can take a smooth, convex, and compact approximation Q_2^R of Q_1^R. Since t is arbitrary, it suffices to show that $Q_2^R \subset \mathbf{B}(Q_2^R : \boldsymbol{\delta})$ for sufficiently small ε.

[2] Precisely, the following explanation is based more on the NEU condition of Abreu, Dutta, and Smith (1994) than on the full dimensionality.

[3] Recently, Guéron, Lamadon and Thomas (2009) show the folk theorem for the specific example in Fudenberg and Maskin (1986) without full dimensionality.

2. To show this, it is important to have $\tilde{v}$ such that

   (a) $\sum_i \lambda_i \frac{1-\delta_i}{\delta_i}\left(v_i^2 - \tilde{v}_i\right) < 0$ and

   (b) there exists $\{\tilde{w}(y)\}$ such that

$$\begin{aligned}
&\tilde{v}_i = (1-\delta)\, g_i(a_i, \alpha_{-i}) + \delta E[\tilde{w}_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
&\tilde{v}_i \geq (1-\delta)\, g_i(a_i, \alpha_{-i}) + \delta E[\tilde{w}_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
&0 \geq \lambda \cdot (\tilde{w}(y) - \tilde{v}) \text{ for all } y \in Y.
\end{aligned}$$

With equal discounting, we have the following:

(a) Since $o \in \mathrm{int}Q^I$, $\max_{v' \in Q_1^I} \lambda \cdot v' - \lambda \cdot \tilde{v} = \max_{v' \in Q_1^I} \lambda \cdot v' - \max_{v' \in Q^I} \lambda \cdot v' < 0$. Hence, for sufficiently good approximation Q_2^I of Q_1^I, $\sum_i \lambda_i \frac{1-\delta}{\delta}\left(v_i^2 - v_i^1\right) = \frac{1-\delta}{\delta}\lambda \cdot \left(v^2 - \tilde{v}\right) < 0$.

(b) Since $\tilde{v} \in \arg\max_{v' \in Q^I} \lambda \cdot v'$, there exists $v(\lambda) \in \arg\max_{v' \in Q^I} \lambda \cdot v'$ such that

$$\begin{aligned}
&v_i(\lambda) = (1-\delta)\, g_i(a_i, \alpha_{-i}) + \delta E[w_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
&v_i(\lambda) \geq (1-\delta)\, g_i(a_i, \alpha_{-i}) + \delta E[w_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
&\lambda \cdot (v(\lambda) - v(\lambda)) = 0, \\
&0 \geq \lambda \cdot (w(y) - v(\lambda)) \text{ for all } y \in Y.
\end{aligned}$$

Therefore, with $\tilde{w}(y) = w(y) + \frac{1}{\delta}(v(\lambda) - \tilde{v})$,

$$\begin{aligned}
&\tilde{v}_i = (1-\delta)\, g_i(a_i, \alpha_{-i}) + \delta E[\tilde{w}_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
&\tilde{v}_i \geq (1-\delta)\, g_i(a_i, \alpha_{-i}) + \delta E[\tilde{w}_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
&0 \geq \lambda \cdot (\tilde{w}(y) - \tilde{v}) \text{ for all } y \in Y.
\end{aligned}$$

The last inequality holds since

$$\lambda \cdot (\tilde{w}(y) - \tilde{v}) = \lambda \cdot (w(y) - v(\lambda) + \frac{1}{\delta}(v(\lambda) - \tilde{v}) + v(\lambda) - \tilde{v}) = \lambda \cdot (w(y) - v(\lambda)).$$

However, with unequal discounting, $\tilde{v} \in \arg\max_{v' \in Q^R} \lambda \cdot v'$ cannot work for the following reasons.

(a) Although $\lambda \cdot (v^2 - \tilde{v}) < 0$ holds, this does not imply $\lambda \cdot \left(\frac{1-\delta_i}{\delta_i} (v_i^2 - \tilde{v}_i) \right)_{i=1}^n < 0$. The latter requires $R\lambda \cdot (v^2 - \tilde{v}) < 0$.

(b) Since $\tilde{v} \in \arg\max_{v' \in Q^R} \lambda \cdot v'$, there exists $v(\lambda) \in \arg\max_{v' \in Q^R} \lambda \cdot v'$ such that

$$\begin{aligned}
&v_i(\lambda) = (1-\delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[w_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
&v_i(\lambda) \geq (1-\delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[w_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
&0 \geq \lambda \cdot (w(y) - v(\lambda)) \text{ for all } y \in Y,
\end{aligned}$$

However, showing the existence of $\{\tilde{w}(y)\}_y$ such that

$$\begin{aligned}
&\tilde{v}_i = (1-\delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[\tilde{w}_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\
&\tilde{v}_i \geq (1-\delta_i) g_i(a_i, \alpha_{-i}) + \delta_i E[\tilde{w}_i(y) : a_i, \alpha_{-i}] \\
&\text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\
&0 \geq \lambda \cdot (\tilde{w}(y) - \tilde{v}) \text{ for all } y \in Y,
\end{aligned} \tag{2}$$

is not straightforward. If we use $\tilde{w}(y)$ with $\tilde{w}_i(y) = w_i(y) + \frac{1}{\delta_i}(v_i(\lambda) - \tilde{v}_i)$, the first two conditions are satisfied while the last one might not be since

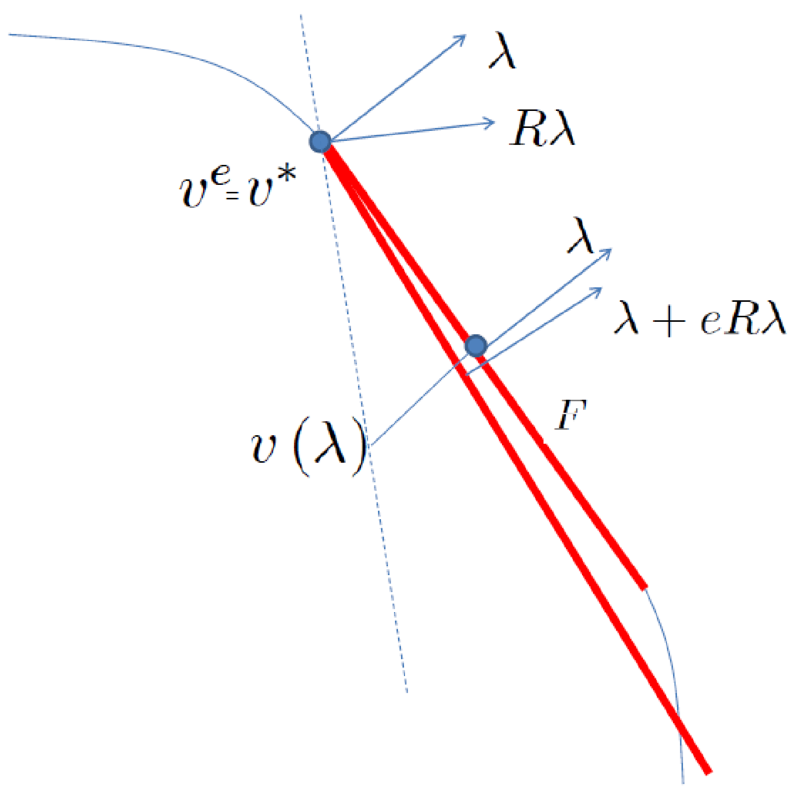
$$\lambda \cdot (\tilde{w}(y) - \tilde{v}) = \lambda \cdot (w(y) - v(\lambda)) + R\lambda \cdot (v(\lambda) - \tilde{v}) \tag{3}$$

and the last term can be large.

Let us consider 2.(b) first. We will construct an approximation Q^e of Q^R such that (2) is satisfied for all λ and $\tilde{v} \in \arg\max_{v' \in Q^e} \lambda \cdot v'$. Let us consider the two-player case for simplicity. If there is a unique maximizer $\arg\max_{v' \in Q^R} \lambda \cdot v'$ for all λ, (2) is satisfied for all λ and $v \in \arg\max_{v' \in Q^e} \lambda \cdot v'$ since otherwise Q^R is not a fixed point. Hence, we concentrate on the case where there is a facet F on Q^R with some normal vector λ. For simplicity, assume that there is a unique maximizer for any $\lambda' \neq \lambda$. Let $v^* = \arg\min_{v' \in F} R\lambda \cdot v'$. Consider the hyperplane passing v^* with the normal vector $\lambda + eR\lambda$ (see Figure below). Let $Q^e = Q^R \cap \{v' \in \mathbb{R}^n : (\lambda + eR\lambda) \cdot v' \leq (\lambda + eR\lambda) \cdot v^*\}$.

For sufficiently small e, for any λ', $v^e \in \arg\max_{v' \in Q^e} \lambda' \cdot v'$, and $v \in \arg\max_{v' \in Q^R} \lambda' \cdot v'$, $R\lambda' \cdot (v - v^e)$ is sufficiently small. Hence, for any λ' and $v^e \in \arg\max_{v' \in Q^e} \lambda' \cdot v'$, since (2) is satisfied for $v(\lambda)$ and we have shown that $R\lambda' \cdot (v(\lambda) - v^e)$ is negligible, v^e (approximately) satisfies (2).

Intuitively, comparing $v^e \in \arg\max_{v' \in Q^e} \lambda' \cdot v'$ and $v(\lambda) \in \arg\max_{v' \in Q^R} \lambda' \cdot v'$, we subtract player 1's continuation payoff from $v(\lambda)$ and add this to player 2's continuation payoff (see Figure below for the case with $\lambda' = \lambda$). Since player 2 is more patient, the effect of this operation increases $\lambda' \cdot v^e$. Hence, (2) is satisfied for v^e.



The remaining thing to show is 2.(a): firstly, we also modify 1. as follows: we take the radial contraction of Q^e with respect to o instead of Q^R. We can take a smooth, convex, and compact approximation Q_2^e of Q_1^e and try to show that $Q_2^e \subset \mathbf{B}(Q_2^e : \boldsymbol{\delta})$. To show this, it is important that for any λ and $v^2 \in \arg\max_{v' \in Q_2^e} \lambda \cdot v'$, there is $v^e \in \arg\max_{v' \in Q^e} \lambda \cdot v'$ such that $R\lambda \cdot \left(v^2 - v^e\right) < 0$ (Note that this corresponds to 2.(a). Note also that 2.(b) is already shown for any v^e). Since Q^e is an approximation of Q^R, the sufficient condition is that there exists $\bar{e} > 0$ such that for all λ and $v \in \max_{v' \in Q^R} \lambda \cdot v'$,

$$R\lambda \cdot (v - o) > 2\bar{e}, \tag{4}$$

which implies that for sufficiently small e, for all λ and $v^e \in \max_{v' \in Q^e} \lambda \cdot v'$,

$$R\lambda \cdot (v^e - o) > \bar{e}. \tag{5}$$

To see why this is sufficient, take any $v^2 \in \arg\max_{v' \in Q_2^e} \lambda \cdot v'$. Since Q_2^e is an approximation of Q_1^e, there exists $v^1 \in \arg\max_{v' \in Q^e} \lambda \cdot v'$ with $R\lambda \cdot \left(v^2 - v^1\right) \approx 0$. Since Q_1^e is inside of Q^e with respect to $R\lambda$, there exists $v^e \in \arg\max_{v' \in Q^R} \lambda \cdot v'$ such that $R\lambda \cdot (v^1 - v^e) < 0$. Hence, $R\lambda \cdot \left(v^2 - v^e\right) < 0$.

Lemma 1 below shows that **Assumption 1** is sufficient for (4). The reason is as follows. While the effect of unequal discounting, represented by R, makes the requirement of (4) stronger, it also "expands" Q^R since there is a room for the intertemporal trade as we explained. These two effects cancel out each other.

The following lemma formalizes the above argument. Note that E' corresponds to Q_2^e with sufficiently small e.

Lemma 1 *If **Assumption 1** is satisfied, there exist $o \in \text{int}Q^I$ and $\bar{e} > 0$ such that, for any compact set E in the interior of Q^R, for any $\varepsilon > 0$, there exists a compact and convex $E' \supset E$ such that*

1. *for all λ and $v \in \max_{v' \in E'} \lambda \cdot v'$,*

$$R\lambda \cdot (v - o) > \bar{e}. \tag{6}$$

and

2. for any λ, $v \in \arg\max_{v' \in E'} \lambda \cdot v'$, *and* $v^R \in \arg\max_{v' \in Q^R} \lambda \cdot v'$,

$$R\lambda \cdot \left(v - v^R\right) < \varepsilon, \tag{7}$$

5 Folk Theorem

In this section, we prove that under the pairwise full rank condition, we can show the folk theorem:

Assumption 2 (Pairwise Full Rank) *For all* $i \neq j$ *and* $\alpha \in A$,

$$\text{rank}\begin{bmatrix} R_i(\alpha) \\ R_j(\alpha) \end{bmatrix} = |A_i| + |A_j| - 1,$$

where $R_i(\alpha)$ *is* $|A_i| \times |Y|$ *matrix with elements* $[R_i(\alpha)]_{a_i y} = \rho(y \mid a_i, \alpha_{-i})$.

Assumption 3 (Individual Full Rank) *For all* i*, there exists an action minmaxing* i *such that, for all* $j \neq i$,

$$\text{rank} R_j(\alpha) = |A_j|.$$

Given **Theorem 1**, it suffices to show that Q^R is equal to the limit of $F(\boldsymbol{\delta})$ under the pairwise full rank condition. Since $F(\boldsymbol{\delta})$ also varies as discounting, we firstly provide the characterization of the limit of $F(\boldsymbol{\delta})$. The characterization is similar to the above except that we do not impose the incentive compatibility condition for the action.

Given a compact W and λ, calculate

$$k^F(\lambda : W : R) = \sup_{v \in IR, \mu \in \triangle(A)} \lambda \cdot v \text{ subject to}$$

$$\begin{cases} v_i = (1-\delta_i) g_i(\mu) + \delta_i w_i \text{ for all } i, \\ 0 \geq \lambda \cdot (w - v), \\ v \in W. \end{cases} \Leftrightarrow \begin{cases} v_i = g_i(\mu) + x_i, \\ 0 \geq R\lambda \cdot x, \\ v \in W, \end{cases}$$

with $x_i \equiv \frac{1-\delta_i}{\delta_i}(w_i - v_i)$ for all i. Define

$$\begin{aligned} H^F(\lambda : W : R) &\equiv \{v : \lambda \cdot v \leq k(\lambda : W : R)\}, \\ \mathcal{B}^F(W : R) &\equiv \bigcap_{\lambda} H^F(\lambda : W : R). \end{aligned}$$

Since $\mathcal{B}^F(\cdot : R)$ is weakly decreasing, $W \subset F^{\square} \Rightarrow \mathcal{B}^F(W : R) \subset F^{\square}$. In addition, $\mathcal{B}^F(W : R)$ is convex, compact, and monotone. Therefore, there exists a largest fixed point of $\mathcal{B}^F(\cdot : R) \subset F^{\square}$ and any fixed point is convex and compact. Let F^R be the largest fixed point of $\mathcal{B}^F(\cdot : R) \subset F^{\square}$.

Then, we can show the following theorem.

Theorem 2 (folk theorem) *1. $F^R \supset F(\boldsymbol{\delta})$ for all $\boldsymbol{\delta}$ with $\delta_i = 1/(1 + r_i\varepsilon)$ for all i.*

*2. If **Assumptions 2** and **3** are satisfied, for all $\boldsymbol{\delta}$ with $\delta_i = 1/(1 + r_i\varepsilon)$ for all i, $\lim_{\varepsilon \to 0} E(\boldsymbol{\delta}) = F^R$.*

The intuition is as follows. Without loss of generality, there exists $a \in A$ that attains $k^F(\lambda : W : R)$. Therefore, it suffices to show that, if there exists a and x such that

$$\begin{cases} v_i = g_i(a) + x_i \geq ir_i, \\ 0 \geq R\lambda \cdot x, \end{cases}$$

then, there exists $\{x(y)\}_{y \in Y}$ such that

$$\begin{aligned} & v_i = g_i(a_i, \alpha_{-i}) + E[x_i(y) : a_i, \alpha_{-i}] \\ & \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ & v_i \geq g_i(a_i, \alpha_{-i}) + E[x_i(y) : a_i, \alpha_{-i}] \\ & \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\ & 0 \geq R\lambda \cdot x(y) \text{ for all } y \in Y. \end{aligned}$$

Assumptions 2 and **3** are sufficient for the above argument since replacing $R\lambda$ with $\tilde{\lambda}$ that is parallel to $R\lambda$ gives us the same problem as FLM. Note that if $\lambda = \pm e_i$ for some i, $R\lambda$ is parallel to λ.

6 Extension and Discussion

6.1 A Path of Convergence

One interpretation of the limit of $\delta_i \to 1$ is that $\boldsymbol{\delta}$ is fixed and the interval between two consecutive repetitions of the stage game goes to 0. As LP pointed out, this approach is equivalent to taking a path of discount factors that converge to 1 while keeping the patience ratio $r_i = \log \delta_i / \log \delta_n$ for all i. While we take a special convergence sequence so that $\delta_i = 1/\left(1 + r_i \varepsilon\right)$ for all i with ε converging to 0 in the previous sections, we can extend the results for any convergence sequence $\{\boldsymbol{\delta}^m\}_{m=1}^{\infty}$ that satisfies $\lim_{m\to\infty} \delta_i = 1$ and $\lim_{m\to\infty} \left(1 - \delta_i^m\right) / \left(1 - \delta_n^m\right) = r_i$ for all i. Since $\log \delta_i / \log \delta_n \approx \left(1 - \delta_i\right) / \left(1 - \delta_n\right)$ in the limit, the sequence in LP is a special case of our generalized convergence sequence.

Theorem 3 *If* ***Assumption 1*** *is satisfied, for all* $\{\boldsymbol{\delta}^m\}_{m=1}^{\infty}$ *with* $\lim_{m\to\infty} \delta_i^{(m)} = 1$ *and* $\lim_{m\to\infty} (1 - \delta_i^{(m)})/(1 - \delta_n^{(m)}) = r_i$ *for all* i, $\lim_{m\to\infty} E(\boldsymbol{\delta}^{(m)}) = Q^R$.

As a corollary, we can extend **Theorem 2**.

Theorem 4 (Folk Theorem) *If* ***Assumptions 1****,* ***2****, and* ***3*** *are satisfied, for all* $\{\boldsymbol{\delta}^m\}_{m=1}^{\infty}$ *with* $\lim_{m\to\infty} \delta_i^{(m)} = 1$ *and* $\lim_{m\to\infty} (1-\delta_i^{(m)})/(1-\delta_n^{(m)}) = r_i$ *for all* i, $\lim_{m\to\infty} E(\boldsymbol{\delta}^{(m)}) = \lim_{m\to\infty} F(\boldsymbol{\delta}^{(m)}) = F^R$.

6.2 Discussions

In this paper, we offer the limit characterization of the PPE and SIR payoffs respectively with unequal discounting. In addition, we show that the pairwise full rank condition is sufficient for the folk theorem. One remaining problem is, as mentioned before, to characterize the set of PPE payoffs when the full dimensionality condition is not satisfied. Fudenberg, Levine, and Takahashi (2007) solve the problem with equal discounting. The basic observation is that if the continuation payoffs are in the subspace of $\mathbb{R}^n$, the payoffs of all the enforceable action profiles by the continuation payoffs should be in the same subspace. This makes it possible to construct a inductive characterization of the PPE payoffs with respect to the dimensionality. With unequal discounting, due to the intertemporal trade, even though the continuation payoffs are in the subspace, it seems hard to derive a restrictions on the payoffs of the enforceable action profiles. We leave this problem for the future research.

7 Appendix

7.1 Proof of Proposition 1

Analogous to APS. The convexity holds from the fact that we allow the correlation μ and do not consider the incentive to take an action.

7.2 Proof of Lemma 1

Since **Assumption 1** is satisfied, we can take $o \in \text{int}Q^I$, where Q^I is the solution for FL problem.

Firstly, we prove (6) for Q^R itself: Suppose not. Then, there exists λ^* and $v^* \in \arg\max_{v \in Q^R} \lambda^* \cdot v$ such that

$$R\lambda^* \cdot v^* \leq R\lambda^* \cdot o. \tag{8}$$

Therefore, since $\max_{v \in Q^I} R\lambda^* \cdot v > R\lambda^* \cdot o$, there exists $\eta > 0$ such that

$$\max_{v \in Q^I} R\lambda^* \cdot v - R\lambda^* \cdot v^* > \eta. \tag{9}$$

"Shift up" v^* by $\varepsilon\lambda^*$: $v^\varepsilon = v^* + \varepsilon\lambda^*$. For sufficiently small ε, $\text{co}\left(\{v^\varepsilon\} \cup Q^R\right)$ satisfies the following: for any λ, there exists $v \in \arg\max_{v' \in \text{co}(\{v^\varepsilon\} \cup Q^R)} \lambda \cdot v'$ with $v \in H(R\lambda)$. To see this, for any λ, either (i) $\arg\max_{\text{co}(\{v^\varepsilon\} \cup Q^R)} \lambda \cdot v \subset Q^R$ or (ii) $v^\varepsilon \in \arg\max_{\text{co}(\{v^\varepsilon\} \cup Q^R)} \lambda \cdot v$. For case (i), it is obvious. For case (ii), for sufficiently small ε, both $\|\lambda - \lambda^*\|$ and $\|v^\varepsilon - v^*\|$ are sufficiently small. Then, we have

$$\begin{aligned} |R\lambda^* \cdot v^* - R\lambda \cdot v^\varepsilon| &< \frac{\eta}{2}, \\ \left|\max_{v' \in Q^I} R\lambda^* \cdot v' - \max_{v' \in Q^I} R\lambda \cdot v'\right| &< \frac{\eta}{2}, \end{aligned}$$

which implies $v^\varepsilon \in H(R\lambda)$. However, this means $\text{co}\left(\{v^\varepsilon\} \cup Q^R\right)$ is a fixed point of $\mathcal{B}(\cdot, R)$, which is the contradiction.

Secondly, we construct E' such that (7) holds. Approximate Q^R by a n-dimensional polygon $\bar{Q}^R$.

For any $\eta > 0$, we can take $\bar{Q}^R$ such that $\bar{Q}^R$ consists of finite $(n-1)$-dimensional facets $\left\{F^k\right\}_{k=1}^K$ and, for any λ and $v \in \arg\max_{v' \in Q^R} \lambda \cdot v'$, there exists $\bar{v} \in \arg\max_{v' \in \bar{Q}^R} \lambda \cdot v'$ such that $\|v - \bar{v}\| < \eta$. Since η is arbitrary, it suffices to construct E' such that, for any λ, $v \in \arg\max_{v' \in E'} \lambda \cdot v'$, and $\bar{v}^R \in \arg\max_{v' \in \bar{Q}^R} \lambda \cdot v$,

$$R\lambda \cdot \left(v - \bar{v}^R\right) < \varepsilon. \tag{10}$$

Let λ^k be the unique tangential vector for F^k. For each k and $e > 0$, let

$$\bar{H}^k \equiv \{v \in \mathbb{R}^n : (\lambda^k + eR\lambda^k) \cdot v \leq \min_{v' \in F^k} (\lambda^k + eR\lambda^k) \cdot v'\}$$

be the hyperplane that is constructed by rotating F^k. Define

$$\bar{Q}^e = \bar{Q}^R \cap \bar{H}^1 \cap \cdots \cap \bar{H}^K.$$

We can guarantee that for all λ^k, $\max_{v' \in \bar{Q}^R \cap \bar{H}^k} (\lambda^k + eR\lambda^k) \cdot v' = \max_{v' \in \bar{Q}^e} (\lambda^k + eR\lambda^k) \cdot v'$ by taking e sufficiently small, that is, no facet $\bar{H}^k$ is completely excluded. Let $\{\bar{F}^k\}_{k=1}^K$ be the set of facets of $\bar{Q}^e$. We show that $E' = \bar{Q}^e$ satisfies (10) for sufficiently small e.

1. Firstly, we show that (10) holds for $\lambda = \lambda^k + eR\lambda^k$.

   Suppose not. Then, for any $\bar{e} > 0$, there exist $e \in (0, \bar{e})$, $v(e) \in \bar{F}^k$, and $v^R(e) \in \arg\max_{v' \in F^k} \left(\lambda^k + eR\lambda^k\right) \cdot v'$ such that

   $$R(\lambda^k + eR\lambda^k) \cdot v(e) \geq R(\lambda^k + eR\lambda^k) \cdot v^R(e) + \varepsilon.$$

   Since $\lambda^k \cdot v'$ is constant for all $v' \in F^k$, without loss of generality, we can pick a fixed $v^R \in \arg\max_{v' \in F^k} R\lambda^k \cdot v'$.

   In addition, since $\bar{Q}^R \supset \bar{Q}^e \supset \bar{F}^k \ni v(e)$, $\max_{v' \in \bar{Q}^R} \lambda^k \cdot v' \geq \lambda^k \cdot v(e)$. At the same time, as e goes to 0, $\left\{\bar{F}^k\right\}_{k=1}^K$ uniformly converges to $\left\{F^k\right\}_{k=1}^K$. Hence, $\lambda^k \cdot v(e) \geq \max_{v' \in \bar{Q}^R} \lambda^k \cdot v' - O(e)$. In summary,

   $$\max_{v' \in \bar{Q}^R} \lambda^k \cdot v' \geq \lambda^k \cdot v(e) \geq \max_{v' \in \bar{Q}^R} \lambda^k \cdot v' - O(e).$$

Note that, since v^R is on the facet F^k,

$$\lambda^k \cdot v^R = \max_{v' \in \bar{Q}^R} \lambda^k \cdot v'.$$

Note also that

$$v(e) \in \bar{Q}^R.$$

Taking subsequence if necessary, the above four inequalities give us the following:

$$\begin{aligned} R\lambda^k \cdot v &\geq R\lambda^k \cdot v^R + \varepsilon, \\ \lambda^k \cdot v &= \max_{v' \in \bar{Q}^R} \lambda^k \cdot v' = \lambda^k \cdot v^R, \\ v &\in \bar{Q}^R, \end{aligned}$$

with $v = \lim_{e \to 0} v(e)$. Since the first two inequalities imply $v \notin \bar{Q}^R$, this is a contradiction.

Since the number of facets is finite, we are done.

2. Note that any vector λ that is not tangential to any facet is expressed as $\sum_{i=1}^{\tilde{n}} \alpha_i (\lambda^{k_i} + eR\lambda^{k_i})$ with $\tilde{n} \leq n$, $\alpha_i > 0$ for all $i = 1, \ldots, \tilde{n}$, $\sum_{i=1}^{\tilde{n}} \alpha_i = 1$, $\lambda^{k_i} + eR\lambda^{k_i}$ being tangential to $\bar{F}^{k_i}$, and $\bar{F}^{k_i} \cap \bar{F}^{k_j} \neq \emptyset$.[4] For sufficiently small e, there exists $v^R \in \arg\max_{v \in \bar{Q}^R} \lambda^{k_i} \cdot v'$ for all $i = 1, \ldots, \tilde{n}$. Consider any $\tilde{v}^R \in \arg\max_{v' \in \bar{Q}^R} \lambda \cdot v'$. Then, since

$$\begin{aligned} 0 &\geq \lambda \cdot (v^R - \tilde{v}^R) \\ &= \sum_{i=1}^{\tilde{n}} \alpha_i \lambda^{k_i} \cdot (v^R - \tilde{v}^R) + eR\tilde{\lambda} \cdot (v^R - \tilde{v}^R) \\ &\geq eR\tilde{\lambda} \cdot (v^R - \tilde{v}^R), \end{aligned}$$

$R\tilde{\lambda} \cdot (v^R - \tilde{v}^R) \leq 0$.

On the other hand, for any $v \in \arg\max_{v' \in \bar{Q}} \lambda \cdot v'$, $v \in \arg\max_{v \in \bar{Q}} \lambda^{k_i} \cdot v'$ for all $i = 1, \ldots, \tilde{n}$.

[4] We identify the vectors with the same direction.

Therefore, from 1., $R\lambda^{k_i} \cdot (v - v^R) \leq \varepsilon$ for all $i = 1, \ldots, \tilde{n}$. Therefore,

$$\begin{aligned} & R\left(\sum\nolimits_{i=1}^{\tilde{n}} \alpha_i \lambda^{k_i}\right) \cdot \left(v - \tilde{v}^R\right) \\ = & \left(\sum\nolimits_{i=1}^{\tilde{n}} \alpha_i R\lambda^{k_i}\right) \cdot \left(v - v^R + v^R - \tilde{v}^R\right) \leq n\varepsilon. \end{aligned}$$

Since n is finite, we are done.

Finally, since (6) is satisfied for Q^R, for sufficiently good approximation $\bar{Q}^R$ for Q^R and sufficiently small e, it is also satisfied for E'.

7.3 Proof of Theorem 1

Proposition 3 *For all $\boldsymbol{\delta}$ with $\delta_i = 1/(1 + r_i\varepsilon)$ for all i, $E(\boldsymbol{\delta}) \subset Q^R$.*

Proof. It suffices to show that $\text{co}(E(\boldsymbol{\delta})) \subset Q^R$. From **Proposition 2**, for all λ, there exists $v \in \arg\max_{v' \in \text{co}(E(\boldsymbol{\delta}))} \lambda \cdot v'$ such that there exist $\alpha \in \mathcal{A}$ and $\{w(y)\}_{y \in Y}$ such that

$$\begin{cases} \{w(y)\}_{y\in Y} \text{ enforces } \langle v, \alpha \rangle, \\ w(y) \in \text{co}(E(\boldsymbol{\delta})) \text{ for all } y \in Y. \end{cases}$$

Since $\lambda \cdot w(y) \leq \lambda \cdot v$ for all $y \in Y$, defining $x(y) = \frac{\delta_n}{1-\delta_n} R^{-1}(w(y) - v)$ for all $y \in Y$,

$$\begin{cases} v_i = g_i(a_i, \alpha_{-i}) + E[x_i(y) : a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ v_i \geq g_i(a_i, \alpha_{-i}) + E[x_i(y) : a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\ 0 \geq R\lambda \cdot x(y) \ \text{ for all } y \in Y. \end{cases}$$

■

For the other direction, the following lemma is helpful.

Lemma 2 *Let $W \subset \mathbb{R}^n$ be convex and compact. If there exist $\boldsymbol{\delta}$ with $\frac{1-\delta_i}{\delta_i} / \frac{1-\delta_n}{\delta_n} = r_i$ for all i and $\eta > 0$ such that $W \cap B_\eta(v) \subset \mathbf{B}(W : \boldsymbol{\delta})$, then, for all $\boldsymbol{\delta}'$ with $\delta_i' \geq \delta_i$ and $\frac{1-\delta_i'}{\delta_i'} / \frac{1-\delta_n'}{\delta_n'} = r_i$ for all i, $W \cap B_\eta(v) \subset \mathbf{B}(W : \boldsymbol{\delta}')$.*

Proof. Since $W \cap B_\eta(v) \subset \mathbf{B}(W : \boldsymbol{\delta})$, for all $v' \in W \cap B_\eta(v)$, there exist α and $\{w(y)\}_{y \in Y}$ such that

$$\begin{cases} \{w(y)\}_{y\in Y} \text{ enforces } \langle v', \alpha \rangle, \\ w(y) \in W \text{ for all } y \in Y. \end{cases}$$

For $\boldsymbol{\delta}'$ with $\delta_i' > \delta_i$ and $\frac{1-\delta_i'}{\delta_i'} / \frac{1-\delta_n'}{\delta_n'} = r_i$ for all i, defining

$$w(y : \boldsymbol{\delta}') \equiv \left(\frac{\delta_i' - \delta_i}{\delta_i'(1-\delta_i)} v_i' + \frac{\delta_i(1-\delta_i')}{\delta_i'(1-\delta_i)} w_i(y) \right)_{i=1}^n ,$$

$\{w(y : \boldsymbol{\delta}')\}_{y \in Y}$ enforces $\langle v', \alpha \rangle$ for $\boldsymbol{\delta}'$. Therefore, it suffices to show that $w(y : \boldsymbol{\delta}') \in W$ for all $y \in Y$.

Since

$$\begin{aligned} & \frac{\delta_i(1-\delta_i')}{\delta_i'(1-\delta_i)} - \frac{\delta_j(1-\delta_j')}{\delta_j'(1-\delta_j)} \\ = \; & \left(\frac{\delta_i}{1-\delta_i} / \frac{\delta_n}{1-\delta_n} \right) \frac{\delta_n}{1-\delta_n} \left(\frac{1-\delta_i'}{\delta_i'} / \frac{1-\delta_n'}{\delta_n'} \right) \frac{1-\delta_n'}{\delta_n'} - \left(\frac{\delta_j}{1-\delta_j} / \frac{\delta_n}{1-\delta_n} \right) \frac{\delta_n}{1-\delta_n} \left(\frac{1-\delta_j'}{\delta_j'} / \frac{1-\delta_n'}{\delta_n'} \right) \frac{1-\delta_n'}{\delta_n'} \\ = \; & 0, \end{aligned}$$

$w(y : \boldsymbol{\delta}')$ is a convex combination of v' and $w(y)$. Since W is convex, $w(y : \boldsymbol{\delta}') \in W$ for all $y \in Y$.

■

Proposition 4 *For all $\boldsymbol{\delta}$ with $\delta_i = 1/(1 + r_i \varepsilon)$ for all i, $\lim_{\varepsilon \to 0} E(\boldsymbol{\delta}) \supset Q^R$.*

Proof. Take any compact $E \subset \text{int} Q^R$. It suffices to show that there exist $\bar{\varepsilon} < 1$ and $\bar{E}$ such that for any $\varepsilon < \bar{\varepsilon}$, $E \subset \bar{E} \subset \mathbf{B}(\bar{E} : \boldsymbol{\delta})$.

From **Lemma 7**, there exist $o \in \text{int} Q^R$ and $\bar{e} > 0$ such that there exists a compact and convex $\hat{E}$ such that

1. there exists $t > 0$ such that $E \subset \hat{E}(t) \equiv \{v \in \mathbb{R}^n : \exists v' \in \hat{E} \text{ such that } v = (1-t)v' + to\}$,

2. for all λ and $v \in \max_{v' \in \hat{E}(t)} \lambda \cdot v'$,

$$R\lambda \cdot (v - o) > \bar{e}.$$

and

3. for any λ, $v \in \arg\max_{v' \in \hat{E}(t)} \lambda \cdot v'$, and $v^R \in \arg\max_{v' \in Q^R} \lambda \cdot v'$,

$$R\lambda \cdot \left(v - v^R\right) < \frac{1}{4} t\bar{e}.$$

Consider $\bar{E} \equiv \bigcup_{v \in \hat{E}(t)} \overline{B_{\frac{1}{4}t\bar{e}}(x)}$. Note that $\bar{E}$ satisfies

1. $E \subset \bar{E}$,

2. for any λ, $v \in \arg\max_{v' \in \bar{E}} \lambda \cdot v'$, there exists $v^R \in Q^R$ such that

$$\begin{cases} v_i^R = g_i(a_i, \alpha_{-i}) + E[x_i(y) \mid a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ v_i^R \geq g_i(a_i, \alpha_{-i}) + E[x_i(y) \mid a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\ 0 \geq R\lambda \cdot x(y) \text{ for all } y \in Y, \\ R\lambda \cdot \left(v - v^R\right) < -\frac{1}{2} t\bar{e}, \end{cases}$$

which implies

$$\begin{cases} v_i = g_i(a_i, \alpha_{-i}) + E[x_i(y) + v_i - v_i^R \mid a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ v_i \geq g_i(a_i, \alpha_{-i}) + E[x_i(y) + v_i - v_i^R \mid a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\ -\frac{1}{2} t\bar{e} \geq R\lambda \cdot (x(y) + v - v^R) \text{ for all } y \in Y. \end{cases}$$

The rest of the proof is analogous to FL. See Mailath and Samuelson (2006) for the details. ■

7.4 Proof of Theorem 2

Firstly, we prove the first argument.

Proposition 5 *$F^R \supset F(\boldsymbol{\delta})$ for all $\boldsymbol{\delta}$ with $\delta_i = 1/(1 + r_i\varepsilon)$ for all i.*

Proof. From **Proposition 1**, for any λ, there exists $v \in \arg\max_{v' \in F(\boldsymbol{\delta})} \lambda \cdot v'$ such that there exist $\mu \in \triangle(A)$ and $w \in F(\boldsymbol{\delta})$ such that $v_i = (1 - \delta_i) g_i(\mu) + \delta_i w_i$ for all i. Since $\lambda \cdot w \leq \lambda \cdot v$, defining

$x = \frac{\delta_n}{1-\delta_n} R^{-1}(w - v)$, v satisfies

$$\begin{cases} v_i = g_i(\mu) + x_i, \\ 0 \geq R\lambda \cdot x, \\ v \in W. \end{cases}$$

■

For the second argument, given above, it suffices to show the following proposition.

Proposition 6 *If **Assumptions 2** and **3** are satisfied, for any W, if $v \in IR \cap W$ satisfies*

$$\begin{cases} v_i = g_i(\mu) + x_i, \\ 0 \geq R\lambda \cdot x, \end{cases}$$

then, there exists $\{x(y)\}_{y \in Y}$ such that

$$\begin{cases} v_i = g_i(a_i, \alpha_{-i}) + E[x_i(y) \mid a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) > 0, \\ v_i \geq g_i(a_i, \alpha_{-i}) + E[x_i(y) \mid a_i, \alpha_{-i}] \\ \text{for all } i \text{ and } a_i \in A_i \text{ such that } \alpha_i(a_i) = 0, \\ 0 \geq R\lambda \cdot x(y) \text{ for all } y \in Y. \end{cases}$$

Proof. We can assume $\mu = a \in \arg\max_{a' \in A} R\lambda \cdot g(a')$ without loss of generality. To see this, consider

$$\begin{aligned} a &\in \arg\max_{a' \in A} R\lambda \cdot g(a') \\ x' &= x - g(a) + g(\mu). \end{aligned}$$

Since $\arg\max_{\mu' \in \triangle(A)} R\lambda \cdot g(\mu') = \arg\max_{a' \in A} R\lambda \cdot g(a')$,

$$\begin{cases} v_i = g_i(a) + x'_i, \\ 0 \geq R\lambda \cdot x', \\ v \in W. \end{cases}$$

Then, the rest of the proof is the same as FLM. ■

7.5 Proof of Theorem 3

Given **Theorem 1**, it suffices to show that Q^R is continuous in R.

Lemma 3 *If **Assumption 1** is satisfied, Q^R is continuous in R.*

Proof. It suffices to show that, for any compact $E \subset \text{int}Q^R$, there exists ε such that, for any R' with $\|R - R'\| < \varepsilon$, there exists $E' \supset E$ such that $\mathcal{B}(E' : R') \subset E'$.

The proof consists of the following two arguments.

1. If **Assumption 1** is satisfied, there exist $o \in \text{int}Q^R$ and $\bar{e} > 0$ such that, for any compact $E \subset \text{int}Q^R$ and $e' > 0$, there exist $E' \supset E$ and $\eta > 0$ such that

   (a) for any λ and $v \in \max_{v' \in E'} \lambda \cdot v'$, $R\lambda \cdot (v - o) > \bar{e}$, and

   (b) for any λ, λ' with $\left\|\lambda - \lambda'\right\| < \eta$, $v \in \arg\max_{v' \in E'} \lambda \cdot v'$, and $v(\lambda') \in \arg\max_{v' \in Q^R} \lambda' \cdot v'$, $R\lambda' \cdot v - R\lambda' \cdot v(\lambda') < e'$.

   From **Lemma 1**, there exists o and $\bar{e} > 0$ such that there exists $E' \supset E$ such that

   (c) for any λ and $v \in \max_{v' \in E'} \lambda \cdot v'$, $R\lambda \cdot (v - o) > \bar{e}$, and

   (d) for any λ, $v \in \arg\max_{v' \in E'} \lambda \cdot v'$, and $v^R \in \arg\max_{v' \in Q^R} \lambda \cdot v$, $R\lambda \cdot \left(v - v^R\right) < e'$.

   Let us define e as in the proof of **Lemma 1**. Then, it suffices to show that $E' = \bar{Q}^e$ satisfies (b) for sufficiently small e.

   Take λ and $v \in \arg\max_{v' \in E'} \lambda \cdot v'$ arbitrarily. Then, there exists $\{\alpha_i, \lambda^{k_i}\}_{i=1}^{\tilde{n}}$ such that $\tilde{n} \leq n$, $\sum_{i=1}^{\tilde{n}} \alpha_i = 1$, $\alpha_i > 0$ for all $i = 1, \ldots, \tilde{n}$, $\lambda = \sum_{i=1}^{\tilde{n}} \alpha_i(\lambda^{k_i} + eR\lambda^{k_i})$ with $\lambda^{k_i} + eR\lambda^{k_i}$ being tangential to $\bar{F}^{k_i}$, and $\bar{F}^{k_i} \cap \bar{F}^{k_j} \neq \emptyset$. As before, any $v^R \in \arg\max_{v' \in \bar{Q}^R} \lambda \cdot v'$ satisfies $v^R \in \arg\max_{v' \in \bar{Q}^R} \lambda^{k_i} \cdot v'$ for all $i = 1, \ldots, \tilde{n}$.

   On the other hand, take any $\lambda' = \sum_{i=1}^{\tilde{n}'} \alpha'_i(\lambda^{k'_i} + eR\lambda^{k'_i})$ such that $\tilde{n}' \leq n$, $\sum_{i=1}^{\tilde{n}'} \alpha'_i = 1$, $\alpha'_i > 0$ for all $i = 1, \ldots, \tilde{n}'$, $\lambda^{k'_i} + eR\lambda^{k'_i}$ being tangential to $\bar{F}^{k'_i}$ and $\bar{F}^{k'_i} \cap \bar{F}^{k'_j} \neq \emptyset$. For any $\bar{\alpha} > 0$, there exists $\bar{\eta} > 0$ such that for any λ' with $\left\|\lambda - \lambda'\right\| < \bar{\eta}$, we can take $\{\lambda^{k_i}\}$ and $\{\lambda^{k'_i}\}$ such that $\alpha_i > \bar{\alpha}$ implies $\lambda^{k_i} \in \{\lambda^{k'_i}\}$. Since the number of facets is finite, we can take $\bar{\eta}$ independently of λ.

Take any $v'^R \in \arg\max_{v' \in \bar{Q}^R} \lambda' \cdot v'$ and consider

$$\begin{aligned} & R\lambda' \cdot (v - v'^R) \\ = \; & R\lambda' \cdot (v - v^R + v^R - v'^R). \end{aligned}$$

As in the proof of **Lemma 1**, $R\lambda \cdot (v - v^R) < n\varepsilon$, which implies $R\lambda' \cdot (v - v^R) < n\varepsilon + r_1 \bar{\eta} \max_{x,y \in F^\square} |x - y|$. In addition,

$$\begin{aligned} & \arg\max_{v' \in \bar{Q}^R} \lambda' \cdot v' \\ = \; & \arg\max_{v' \in \bar{Q}^R} (\sum_{i=1}^{\tilde{n}'} \alpha_i' \lambda^{k_i'} \cdot v' + e \sum_{i=1}^{\tilde{n}'} \alpha_i' R \lambda^{k_i'} \cdot v') \end{aligned}$$

and

$$\sum_{i=1}^{\tilde{n}'} \alpha_i' \lambda^{k_i'} \cdot v^R \geq \max_{v' \in \bar{Q}^R} \sum_{i=1}^{\tilde{n}'} \alpha_i' \lambda^{k_i'} \cdot v' - n\bar{\alpha} \max_{x,y \in F^\square} |x - y|$$

imply

$$R\lambda' \cdot (v^R - v'^R) \leq \frac{n\bar{\alpha}}{e} \max_{x,y \in F^\square} |x - y|$$

by the following reason. Suppose not. Then, $v^R \neq v'^R$ and $R\lambda' \cdot (v^R - v'^R) > \frac{n\bar{\alpha}}{e} \max_{x,y \in F^\square} |x - y|$. However,

$$\begin{aligned} & \lambda' \cdot v^R - \lambda' \cdot v'^R \\ = \; & \sum_{i=1}^{\tilde{n}'} \alpha_i' \lambda^{k_i'} \cdot (v^R - v'^R) + eR\lambda' \cdot (v^R - v'^R) \\ > \; & 0, \end{aligned}$$

which is a contradiction.

Since $\bar{\alpha}$ is independent of e, by taking sufficiently small $\bar{\alpha}$, we are done with (a).

2. From **Lemma 1**, there exist $o \in \text{int}Q^R$ and $\bar{e} > 0$ such that $R\lambda \cdot (v - o) > \bar{e}$ for all λ and $v \in \max_{v' \in Q^R} \lambda \cdot v'$.

   Take any compact $E \subset \text{int}Q^R$. Then, there exists a compact $\bar{E}$ with $E \subset \text{int}\bar{E}$ and $\bar{E} \subset \text{int}Q^R$. Then, there exists $t > 0$ such that, for any $\bar{E}' \supset \bar{E}$, $\bar{E}'(t) = \{v \in \mathbb{R}^n : \exists v' \in \bar{E}'$ such that $v = (1 - t)v' + to\}$ satisfies $E \subset E'(t)$.

From 1, there exist $\bar{E}' \supset \bar{E}$ and $\eta > 0$ such that

(a) for any λ and $v \in \max_{v' \in \bar{E}'} \lambda \cdot v'$, $R\lambda \cdot (v - o) > \bar{e}$, and

(b) for any λ, λ' with $\left\|\lambda - \lambda'\right\| < \eta$, $v \in \arg\max_{v' \in \bar{E}'} \lambda \cdot v'$, and $v(\lambda') \in \arg\max_{v' \in Q^R} \lambda' \cdot v'$, $R\lambda' \cdot v - R\lambda' \cdot v(\lambda') < t\frac{1}{2}\bar{e}$.

Then, there exists $\varepsilon > 0$ such that, for all R' with $\|R - R'\| < \varepsilon$, $\lambda' = (R')^{-1} R\lambda$ satisfies $\left\|\lambda - \lambda'\right\| < \eta$. Then, for any λ and $v(t) \in \arg\max_{v' \in \bar{E}'(t)} \lambda \cdot v'$, there exists $v \in \arg\max_{v' \in \bar{E}'} \lambda \cdot v'$ such that $R\lambda \cdot (v - v(t)) \geq t\bar{e}$. For v, there exists $v^R \in H(R\lambda')$ such that $R'\lambda \cdot v - R'\lambda \cdot v^R = R\lambda' \cdot v - R\lambda' \cdot v^R < t\bar{e}$, which implies $v(t) \in H(R'\lambda)$. Therefore, $\bar{E}'(t)$ satisfies $\mathcal{B}^F\left(\bar{E}'(t) : R\right)$ and $E \subset \bar{E}'(t)$.

■

References

[1] Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58(5), 1041-1063.

[2] Chen, B. (2008): "On Effective Minimax Payoffs and Unequal Discounting," *Economics Letters*, 100(1), 105-107.

[3] Fudenberg, D., and D. K. Levine (1994): "Efficiency and Observability with Long-Run and Short-Run Players," *Journal of Economic Theory*, 62(1), 103-135.

[4] Fudenberg, D., D. K. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62(5), 997-1039.

[5] Fudenberg, D., D. K. Levine, and S. Takahashi (2007): "Perfect Public Equilibrium when Players are Patient," *Games and Economic Behavior*, 61(1), 27-49.

[6] Fudenberg, D. and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533-554.

[7] Gueron, Y., T. Lamadon, and C. Thomas (2009) "On the Folk Theorem with One-Dimensional Payoffs and Different Discount Factors," mimeo

[8] Haag, M., and R. Lagunoff (2007): "On the Size and Structure of Group Cooperation," *Journal of Economic Theory*, 135(1), 68-89.

[9] Harrington, J. (1989): "Collusion among Asymmetric Firms: The Case of Different Discount Factors," *International Journal of Industrial Organization*, 7(2), 289-307.

[10] Lehrer, E., and A. Pauzner (1999): "Repeated Games with Differential Time Preferences," *Econometrica*, 67(2), 393-412.

[11] Salonen H. and H. Vartiainen (2008): "Valuating Payoff Streams under Unequal Discount Factors," *Economic Letters*, 99(3), 595-598.