The folk theorem for repeated games with time-dependent discounting *

Xiaoxi LI †

August 16, 2019

Abstract

This paper defines a general framework to study infinitely repeated games with time-dependent discounting, in which we distinguish and discuss both time-consistent and time-inconsistent preferences. To study the long-term properties of repeated games, we introduce an asymptotic condition to characterize the fact that players become more and more patient, that is, the discount factors at all stages uniformly converge to 1. Two types of folk theorems are proven under perfect observations of past actions and without the public randomization assumption: the asymptotic one, i.e. the equilibrium payoff set converges to the individual rational set as players become patient, and the uniform one, i.e. any payoff in the individual rational set is sustained by a single strategy profile which is an approximate subgame perfect Nash equilibrium in all games

*This research is supported by the National Natrual Science Foundation of China (ref.11701432) and the Central Universities's Research Funding of China (ref.413000052). Part of the work was done when the author was a teaching and research fellow (ATER) at THEMA of Université Cergy-Pontoise, France during the academic year 2015-2016. The author is grateful to Eilon Solan, Jong Jae Lee, Xiangyu Qu, Joel Sobel, and Tristan Tomala for discussions and helpful comments.

†Economics and Management School of Wuhan University, Luojia Hill, Wuhan China, 430072. Email: xxleewhu@gmail.com

with sufficiently patient discount factors. As corollaries, our results of time-inconsistency imply the corresponding folk theorems with the quasi-hyperbolic discounting.

JEL Classifications: C72; C73; D90

Keywords: Repeated games; folk theorem; subgame perfect Nash equilibrium; uniform equilibrium; time-dependent discounting ; time-inconsistency; quasi-hyperbolic discounting

1 Introduction

Repeated games is a basic model for the study of dynamic strategic interactions. One classical problem in the literature of repeated games asks whether, and how, long-run interaction among players enables cooperation, leading to more efficient equilibrium outcomes unable to be realized in the short run. In particular, the folk theorem states that any *feasible and individual rational* payoff vector of the stage game can be sustained by a subgame perfect Nash equilibrium in the repeated game as the players become *sufficiently patient*.

The early development of the theory attempted to obtain various versions of folk theorems under different evaluation criteria: the *undiscounted* model in Aumann and Shapley [2] and in Rubinstein [23]; the *discounted* model in Sorin [26], and in Fudenberg and Maskin [9, 11]; the *finite-stage average* model in Benoit and Krishna [3, 4], and in Gossner [12]. These positive results ensured in somehow the robustness of the folk theorem with respect to the way the payoff streams being evaluated[1], essentially how "sufficiently patient" being characterized. Later on, the main research

[1]We notice that there are different minor conditions on the stage game for the folk theorems to hold, for example, the finite-stage average model assumes non-uniqueness of the Nash equilibrium payoffs, the discounted model assumes full dimension of the feasible and individual rational payoff set.

theme shifted to discounted repeated games with imperfect monitoring of the past actions (see Mailath and Samulson [17] for a comprehensive review).

Our paper focuses on the discounted model. By introducing some generalized asymptotic condition for characterizing players being patient, we prove the folk theorem when the discount factor is time dependent. In dynamic decision problems, discounting helps sum or compare the future payoffs to the current. The discounting model[2] is appropriate and especially useful when the agent has a separable preference for intertemporal choices, which therefore can be represented by an additive utility function (cf. Lowenstein and Prelec [16]). The traditional discounting model assumes the *stationarity* with respect to time, so the time preference can be characterized by a single discount factor. Numerous experimental results (documented in Frederick, Loewenstein and O'Donoghue [8]) suggested that the decision maker's time preference is *unstationary* in a context of dynamic decisions, and sometimes exhibited particularly a *present-biased* property. This phenomenon of *time inconsistency* is captured in theory by the so-called *quasi-hyperbolic discounting* (cf. Laibson[15]).

To study the folk theorem under time-dependent discounting, we distinguish and discuss two different models, time-consistent and time-inconsistent preferences. Under time consistency, players' preference is characterized by a single sequence of discount factors $(\delta_t)_{t\geq 1}$, where δ_t stands for the discount rate of the "calender" stage-t payoff into stage $t-1$, no matter in which stage is the current play. Time inconsistency arises when the player discounts the same amount of payoff at a fixed stage differently as the play proceeds. To model this, we assume a class of sequences $\{(\delta_{T,t})_{t\geq 1} : T \geq 1\}$,

[2]The justification for discounting could be multiple, for example, the *time preference* (a current unit of consumption/outcome is always preferred over a delayed one), the *risk aversion* (there is always uncertainty over future, say a motality rate), or the *interest rate* (investment opportunity with certain positive reward). When no confusion is made, we shall use also the term "time preference" in the later context for a general utility function with discounting that may refer to any of the justifications. The interested reader is referred to Frederick, Loewenstein and O'Donoghue [8] for a comprehensive review on the use of discounting models in economics.

where $\delta_{T,t}$ refers to the discount rate of the "calender" stage-t payoff into stage $t-1$, when the play is currently in stage T. That is, there are multi-selves each one acting at a stage, who makes a decision by evaluating the future payoffs in her/his own way. A usually seen case presenting some sort of stationarity[3] (incorporating the quasi-hyperbolic discounting) is to assume that $\delta_{T,t}$ depends on T and t only through $t-T$.

We adopt the solution concept of *Strotz-Pollak* equilibrium[4] for the study of subgame perfect Nash equilibrium in repeated games with time-inconsistent preferences. for which checking the equilibrium optimality involves the one-shot deviation at a single stage only. The assumption underlying this definition is that under time inconsistency, the multi-selves at different stages will each act for her/his "own" expected payoff only, so when deviating in the current stage, her/his future selves will not coordinate with her/him, hence sticking to the equilibrium plan.

This equilibrium criterion of checking one-shot deviation only for time-inconsistent preferences shares the same spirit of that for the subgame perfect Nash equilibrium under the time-consistent preferences. In viewing this, we are able to unify the proofs of the two parts, by first putting the main focus on the proof for the time-consistent folk theorem, pointing out the main differences, and finally completing the proof for that of time inconsistency.

To characterize the fact that players become more and more patient in this general framework of time-dependent discounting, we first introduce an asymptotic condition on a sequence of time-dependent discount factors $(\delta_t^k)_{k\geq 1}$ (*resp.* $\{(\delta_{T,t}^k) : T \geq 1\}_{k\geq 1}$ for time inconsistency), which is, the discount factors at all stages δ_t^k (*resp.* $\delta_{T,t}^k$) converge

[3]Obara and Park [21] presented a general model of time preferences, which does not assume time-additivity as we do here. Nevertheless, they still assumed such "stationarity" for time-inconsistent preferences.

[4]Initially discussed by Strotz [25] for dynamic decision problems, formalized by Peleg and Yaari [22], and introduced and studied for repeated games by Chade et al. [7]. In Obara and Park [20], it is called the *agent subgame perfect Nash equilibrium*.

uniformly in t (*resp.* in both t and T) to 1 as k tends to infinity. In the particular case of stationary discounting, it corresponds to the condition that the discount factor "$\delta \to 1$". Our main contribution is to establish an asymptotic folk theorem with general discounting (in versions of both time consistency and inconsistency) under such a patience condition: the subgame perfect Nash equilibrium (*resp.* Strotz-Pollak equilibrium) payoff set converges to the feasible and individual rational set as players become sufficiently patient. Within the same framework, we establish also a perfect folk theorem for the notion of uniform equilibrium (which is extensively studied in Mertens et al.[19]). Basically, a strategy profile is a subgame perfect uniform equilibrium if it is an approximate subgame perfect Nash equilibrium in all games with sufficiently patient discount factors. This equilibrium concept, similar while slightly stronger equilibrium notion than the "liminf" or "limsup" criterion used in the undiscounted model ([2], [23]), sacrifices the zero-optimality for the robustness of a single equilibrium strategy profile.

Our equilibrium profile's construction follows the similar line as in [9, 11], which is defined by three phases, *main path*, *punishing*, and *rewarding*. The detection of a deviation within any phase will lead to an immediate collective punishment of the deviator. To ensure the subgame perfection of the equilibrium, the punishment will not last forever. If the punishing phase of finite duration ends with no deviation, the play moves to the rewarding phase. Within this third phase, all players other than the previous deviator are "rewarded" in the sense that on the equilibrium path each of them receives a strictly positive bonus above the specified level. When considering time-dependent discounting, the *stationarity* of the time preference disappears, in particular, starting from a different stage T, players' continuation time preference over future will be different[5]. This will bring us two main challenges when adopting

[5]The reasons are different for time-consistent and time-inconsistent preferences. With time-inconsistency, a different self is acting at her/his stage with a different sequence of discount factors. As for time-consistent preference, even though the sequence of discount factors is the same across stages, being truncated at different stages, the distribution of the sequence over the remaining stages

the proof strategy of [9, 11] that we are going to briefly discuss below.

The punishing phase, in between the main and the rewarding, has a length properly specified in order to incentivize both the deviator, long enough to deter any deviation, and the punishers, not too long so the incurred loss can be compensated later on. In the stationary discounting case, this length is fixed irrespective of the stage on which the deviation occurred, due to the fact that players use a stationary evaluation (normalized weight of each stage) for the future payoff flow. This is no longer the case under time-dependent discounting, so one main difference from that of [9, 11] is the length of the punishment phase is different for deviation at different stages. This makes the equilibrium strategy's optimality checking much more involved.

The same as in the stationary discounting case, one major difficulty for the study of subgame perfect Nash equilibrium in repeated games stems from the dispensation of the *public randomization*[6]. For a feasible payoff vector x to be realized, one has to construct an infinite sequence of pure actions[7] leading to it. The problem is that players may deviate at any time during the play, a further requirement asks that, when the game is truncated at any stage, the average discounted payoff in the remaining game is sufficiently close to x when players become patient. Under stationary discounting, the proof of the aforementioned condition makes the main content of [11] in dispensation of the public randomization assumed in [9]. Moreover, for the construction of the pure action sequence, [26] and [11] employed two different algorithms. In the context of general discounting, our proof needs a perturbed thus

is in general different.

[6]To assume the availability of public randomization is equivalent to assume the observability of the mixed strategy being played at each stage.

[7]Being differently, mixed strategies instead of infinite sequences of pure actions are used in the punishing phase to minmaxize the deviator. To understand this, notice that the punishing phase should be of finite length for the reason of subgame perfection. On the other hand, the punishing phase is ended with a rewarding phase, the payoff of which makes punishers no incentive to deviate even within the support of the minmax mixed strategy.

generalized version of the result (cf. Proposition 4.3). It seems that the algorithm in [11] does not extend, our construction is motivated by [26] and takes a much more intricate form.

There is recently a growing interest paid to repeated games with a general discounting structure, notable[8] Obara and Park [20, 21]. Within the context of time incosistent discounting, there is a special focus paid to the study of repeated games with quasi-hyperbolic discounting or present-biased preferences. Chade et al. [7] is a study of the comparative statics of the equilibrium payoff set with respect to the discount factor, and Bernergard [5] obtained a folk-theorem type result. We notice that Obara and Park [20] provided a folk theorem for future-biased preferences where players use symmetric strategies in a symmetric stage game. We establish folk theorems for time-inconsistent discounting of preferences that take a sufficiently general form. While the previous strand of literature focused more on time inconsistency, both Bernheim and Dasgupta [6] and Arribas and Urbano [1] take time consistency and assume some general structure of discounting. In both of the two papers, only a particular class of repeated games is studied: the stage game as a generalized prisoners' dilemma, i.e. having a unique dominant Nash equilibrium whose payoff is strictly dominated. We put no restrictions on the stage games except for the usual full-dimension condition of the feasible and individual rational payoff set. We emphasize that almost all the above mentioned folk-theorem results are obtained by assuming the availability of a public randomization device or the existence of pure equilibrium/minmax strategies of the stage game, tackeling this technically difficulty is a main theme of our paper.

The organization of the paper is as follows. In section 2, we introduce the model of time-dependent discounting and focus on time consistency. Section 3 presents the formal model of the repeated game. Our main results for time-consistent discounting

[8]See Kochov and Song [13] as well for the time-dependent discount factor being endogenously determined by the previous actions played.

are stated in section 4. Section 5 contains several preliminary results, and section 6 is devoted to the proofs of the main results. Section 7 discusses the model of time-inconsistent discounting and presents the results.

2 The model

In this section we describe the model of repeated games with time-dependent discounting. The focus is on the time consistency for the moment, and the time-inconsistent discounting is discussed in Section 6.

2.1 The time-dependent discounting

Let $(\delta_t)_{t\geq 1}$ be time-dependent discount factors, where $\delta_1 = 1$ and $\delta_t \in [0,1]$ is the discount rate at stage $t-1$ for the payoff at stage $t \geq 2$, i.e. one unit of payoff at stage t is equivalent to δ_t unit of payoff at stage $t-1$. We define the discount rate δ_t (hence stage-t evaluation θ_t) referring to the "physical" time t, i.e. δ_t depends only on the stage on which the payoff received but not on the stage in which the player currently locates, so there is no time inconsistency for such preference. We leave the discussion on time-inconsistent preferences to be in Section 6.

Computing in a recursive way, one unit payoff at stage t is then discounted to stage 1 as $\prod_{1\leq s\leq t}\delta_s$. Let $(g_t)_{t\geq 1}$ be a sequence of stage payoff in a dynamic decision problem, then the total $(\delta_t)_{t\geq 1}$-discounted payoff is $\sum_{t\geq 1}^{+\infty}\big(\prod_{1\leq s\leq t}\delta_s\big)g_t$. Assuming $\sum_{t\geq 1}^{+\infty}\big(\prod_{1\leq s\leq t}\delta_s\big) < +\infty$ (implied by the condition $\sup_{t\geq 1}\delta_t < 1$), the normalized total discounted payoff is then

$$\Big(\frac{1}{\sum_{t'\geq 1}^{+\infty}\big(\prod_{1\leq s\leq t'}\delta_s\big)}\Big)\sum_{t\geq 1}^{+\infty}\big(\prod_{1\leq s\leq t}\delta_s\big)g_t.$$

We introduce a compact way to write the above expression. For given discount factors $(\delta_t)_{t\geq 1}$, we define a sequence $(\theta_t)_{t\geq 1}$ in $[0,1]$ with $\theta_t = \frac{\prod_{1\leq s\leq t}\delta_s}{\sum_{t'\geq 1}^{+\infty}(\prod_{1\leq s\leq t'}\delta_s)}$, which is the weight for the stage-t payoff. Then the normalized discounted payoff for $(g_t)_{t\geq 1}$

is $\sum_{t\geq 1}^{+\infty} \theta_t g_t$. One observes that:

(P1) θ_t is decreasing in t;

(P2) $\sum_{t\geq 1}^{+\infty} \theta_t = 1$.

Definition 2.1 *A sequence $(\theta_t)_{t\geq 1}$ in $[0,1]$ satisfying the above properties (P1) and (P2) is called an **evaluation**.*

On the other hand, given an evaluation $(\theta_t)_{t\geq 1}$, it is easy to establish the associated time-dependent discount factors $(\delta_t)_{t\geq 1}$ equivalent to it, i.e. one puts[9] $\delta_1 = 1$ and $\delta_t = \frac{\theta_t}{\theta_{t-1}}, \forall t \geq 2$. Therefore, to study repeated games with time-dependent discount factors, it is sufficient for us to focus on the *evaluations*.

Example 1 *In the classical models,*

(1) the stationary discounting with a factor δ corresponds to the evaluation $\theta_t = (1-\delta)\delta^{t-1}$;

(2) the T-stage average corresponds to the evaluation $\theta_t = \frac{1}{T}\mathbb{1}_{[1,T]}(t)$.

Let $(\theta^k)_{k\geq 1}$ be a sequence of evaluations with each $\theta^k = (\theta^k_t)_{t\geq 1}$. To study the long-term properties of the repeated game, a first task is to define an asymptotic condition on $(\theta^k)_{k\geq 1}$ for players to be *patient*. For the aim of incorporating the traditional asymptotic condition $''\delta \to 1''$ as a special case and also due to technical reasons, we introduce the following

Definition 2.2 *$(\theta^k)_{k\geq 1}$ is a sequence of **patient** evaluations if:*

$$\forall t \in \mathbb{N}^*,\ \lim_{k\to\infty} \frac{\theta^k_t}{\theta^k_{t+1}} = 1, \quad \textit{and the convergence is uniform in } t.$$

Remark. (1). Since $\sum_{t\geq 1} \theta^k_t = 1$ for all $k \geq 1$, $(\theta^k)_{k\geq 1}$ being patient implies that $\sup_{t\geq 1} \theta^k_t = \theta^k_1$ vanishes as k tends to infinity.

(2). The asymptotic condition for a sequence of evaluations being patient incorporates $''\delta \to 1''$ for the δ-discounted model, and $''\beta \to 1, \delta \to 1''$ for the (β,δ)-preference.

[9] We set $\delta_t = 0$ whenever $\theta_t = 0$.

Nevertheless, our condition does not include the asymptotic condition $''T \to +\infty''$ for the T-stage average model. To incorporate all of them, one may define other weaker asymptotic condition of patience, for example, $\sum_{t\geq 1}^{+\infty} |\theta_t^k - \theta_{t+1}^k| \longrightarrow_{k\to+\infty} 0$ as in Renault [18]. However, this does not necessarily give the technical conditions needed for the proof. One open question is then to look for another asymptotic condition of patience for which: a) $''\delta \to 1''$, and $''T \to +\infty''$ are all incorporated as special cases, and b) the folk theorem is obtained[10].

(3). Let $(\delta_t^k)_{k\geq 1, t\geq 1}$ be the sequence of time-dependent discount factors associated with $(\theta^k)_{k\geq 1}$. By definition, the sequence of evaluations $(\theta^k)_{k\geq 1}$ being patient is equivalent to: $\forall t \geq 1$, the sequence $(\delta_t^k)_{k\geq 1}$ converges uniformly (in t) to 1 as k tends to infinity.

2.2 The repeated games

There is a set of players I. Each player $i \in I$ has a finite action set A^i and we denote $A = \prod_{i\in I} A^i$. Let $g^i : A \to \mathbb{R}$ be the stage payoff function for player i, and we write $g = (g_i)_{i\in I}$ for the vector mapping defined from A to $\mathbb{R}^I$. We assume *perfect observation* and the repeated game denoted by Γ is played as follows. At each stage $t \geq 1$, after observing the actions $(a_1, ..., a_{t-1})$ being played at previous stages, the players independently and simultaneously choose an action in their own set of actions.

For a set S, let $\Delta(S)$ be the set of probability distributions over S. A *behavior strategy* for player i is denoted as $\sigma^i = (\sigma_t^i)_{t\geq 1}$, where for each $t \geq 1$, σ_t^i is a function from the set of t-stage histories $\mathcal{H}_t := \left(\prod_{j\in I} A^j\right)^{t-1}$ to $\Delta(A^i)$ (by convention, $\mathcal{H}_1$ is the singleton set $\{\emptyset\}$). Let Σ^i be the set of player i's behavior strategies and $\Sigma = \prod_{i\in I} \Sigma^i$ is the joint set. Any strategy profile $\sigma = (\sigma^i)_{i\in I} \in \Sigma$ naturally defines a probability

[10] We know from Benoit and Krishna [3] that the folk theorem under T-stage average does not hold unless we assume for distinct stage-game Nash equilibrium payoffs. We may expect that with or without distinct stage-game Nash equilibrium payoffs, assuming some different general patience conditions is needed.

distribution over $\mathcal{H}_\infty := \left(\prod_{j\in I} A^j\right)^\infty$ the set of infinite histories, which is endowed the product sigma-field $\sigma(\{\mathcal{H}_t, t \geq 1\})$. Let $\mathbb{P}_\sigma[\cdot]$ be this probability distribution and $\mathbb{E}_\sigma[\cdot]$ is the corresponding expectation operator.

For a joint profile $a \in A$, we may write $a = (a^i, a^{-i}) \in A^i \times A^{-i}$ $\big($*resp.* $\sigma = (\sigma^i, \sigma^{-i}) \in (\Sigma^i, \Sigma^{-i})\big)$ with $A^{-i} := \prod_{j\neq i} A^j$ $\big($*resp.* $\Sigma^{-i} := \prod_{j\neq i} \Sigma^j\big)$ to stress the i-th component.

Let $\theta = (\theta_t)_{t\geq 1}$ be an *evaluation*, which is by definition a decreasing sequence in $[0,1]$ with the sum $\sum_{t\geq 1}^{+\infty} \theta_t$ equal to 1. For any history $h_\infty = (a_t)_{t\geq 1} \in \mathcal{H}_\infty$, its evaluation under θ is $\gamma_\theta(h_\infty) = \sum_{t\geq 1}^{+\infty} \theta_t g(a_t)$. The θ-evaluated payoff associated with any behavior strategy profile $\sigma \in \Sigma$ is defined as $\gamma_\theta(\sigma) = \mathbb{E}_\sigma[\sum_{t\geq 1} \theta_t g(a_t)]$. Γ_θ refers to the repeated game with the specified evaluation θ. Each player i in Γ_θ aims at maximizing his θ-evaluated payoff $\gamma_\theta^i(\sigma)$ by choosing his strategy σ^i in Σ^i while taking into account that the other players use some strategies σ^{-i} in Σ^{-i}.

Definition 2.3 *A strategy profile $\sigma \in \Sigma$ is a **Nash equilibrium** (resp. ε-**Nash equilibrium**, $\varepsilon > 0$) of the repeated game Γ_θ if $\forall i \in I$, $\forall \tau^i \in \Sigma^i$,*

$$\gamma_\theta^i(\sigma^i, \sigma^{-i}) \geq \gamma_\theta^i(\tau^i, \sigma^{-i}) \qquad \Big(\textit{resp.}\ \ \gamma_\theta^i(\sigma^i, \sigma^{-i}) \geq \gamma_\theta^i(\tau^i, \sigma^{-i}) - \varepsilon\Big).$$

Notation. For any $\theta = (\theta_t)_{t\geq 1}$ an evaluation and T a positive integer, we write $\theta[T] = \big(\theta_t[T]\big)_{t\geq 1}$ for the evaluation as a *weighted continuation of θ from stage T on*, i.e. $\theta_t[T] = \frac{\theta_{T-1+t}}{\sum_{s\geq T} \theta_s}, \forall t \geq 1$.

After a T-stage history $h_T = (a_1, ..., a_{T-1}) \in \mathcal{H}_T$, the *subgame* $\Gamma_{\theta[T]}(h_T)$ is a repeated game as a continuation of Γ_θ with the evaluation $\theta[T]$. With the fixed history h_T, any behavior strategy $\sigma = (\sigma_t)_{t\geq 1}$ in the original game Γ_θ naturally induces a behavior strategy $\sigma_{|h_T} = (\sigma_{t[h_T})_{t\geq 1}$ in the continuation game $\Gamma_{\theta[T]}(h_T)$ as follows:

$$\forall h_s \in \mathcal{H}_s, \quad (\sigma|_{h_T})_s(h_s) = \sigma_{T+s-1}(h_T h_s),$$

where $h_T h_s$ is a $(T + t - 1)$-stage history as a *concatenation* of h_T with h_s, i.e. $h_T h_s = (a_1, ..., a_{T-1}, b_1, ..., b_{s-1})$ for $h_T = (a_1, ..., a_{T-1})$ and $h_s = (b_1, ..., b_{s-1})$.

Definition 2.4 *A strategy profile $\sigma \in \Sigma$ is a* ***subgame perfect Nash equilibrium*** *(resp.* ***subgame perfect ε-Nash equilibrium****, $\varepsilon > 0$) of the repeated game Γ_θ if for all history $h_T \in \mathcal{H}_T, T \geq 1$, $\sigma_{|h_T}$ is a Nash equilibrium (resp. ε-Nash equilibrium) of the subgame $\Gamma_{\theta[T]}(h_T)$.*

3 The main results

We introduce the following notations for the statement of a folk theorem.

Notation. We denote $v = (v^i)_{i\in I}$ where

$$v^i = \min_{s\in\prod_{j\neq i}\Delta(A^j)} \max_{a^i\in A^i} \sum_{a^{-i}\in A^{-i}} s(a^{-i})g^i(a^i, a^{-i}), \quad \forall i \in I.$$

v^i is the *minmax levels*, i.e. the worst payoff level for player i that can be forced by all players other than i without correlation. Let $S = co\{g(a)|a \in A\}$ be the set of feasible payoff vectors, where co is the convex hull operator. $\mathcal{F} := \{x \in S|x^i \geq v^i, \forall i \in I\}$ is the set of *feasible* and *individual rational* payoff vectors, and $int\mathcal{F}$ is the *interior* set of $\mathcal{F}$. For each $i \in I$, let $\hat{s}^i \in \prod_{j\neq i}\Delta(A^j)$ be some product strategy profile such that

$$\sum_{a^{-i}\in A^{-i}} \hat{s}^i(a^{-i})g^i(a^i, a^{-i}) \leq v^i, \ \forall a^i \in A^i.$$

We denote by E_θ the subgame perfect Nash equilibrium payoff set of the repeated game Γ_θ, i.e. for any $x \in E_\theta$, there exists a subgame perfect Nash equilibrium $\sigma \in \Sigma$ with $\gamma_\theta(\sigma) = x$.

Theorem 3.1 ***(Asymptotic Perfect Folk Theorem)*** *Let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations. Suppose that in a repeated game Γ, $int\mathcal{F}$ is nonempty and fully dimensional, then $E_{\theta^k} \longrightarrow_{k\to\infty} \mathcal{F}$, where the convergence is with respect to Hausdoff distance.*

Let $\Lambda := (\delta_t)_{t\geq 1}$ be time-dependent discount factors. We write $\theta(\Lambda)$ for the equivalent evaluation associated with Λ. When no confusion is caused, we may write shortly

$\gamma_\Lambda(\sigma)$ $\big(resp.\ \gamma_{\Lambda[T]}(\sigma)\big)$ for the payoff $\gamma_{\theta(\Lambda)}(\sigma)$ $\big(resp.\ \gamma_{\theta(\Lambda)[T]}(\sigma)\big)$. In the same way, we write Γ_Λ (*resp.* E_Λ) for the repeated game $\Gamma_{\theta(\Lambda)}$ (*resp.* for the perfect equilibrium payoff set $E_{\theta(\Lambda)}$).

The above result can be restated in the following form.

Theorem 3.2 ***(Asymptotic Perfect Folk Theorem restated)*** *Suppose that in a repeated game Γ, $int\mathcal{F}$ is nonempty and fully dimensional. Let x be any vector in $int\mathcal{F}$. For any sufficiently small ε, there exists some $\delta_0 \in (0,1)$ such that for all $\Lambda = (\delta_t)_{t\geq 1}$ with $\delta_t \in [\delta_0, 1), \forall t \geq 1$, we have: there is some $y \in E_\Lambda$ satisfying $\|x - y\|_\infty \leq \varepsilon$.*

The above results concern the convergence of equilibrium payoff set of given discount factors (evaluation) as players become patient. In particular, the equilibrium strategy to sustain each payoff vector in the individual rational set depends on the specified discount factors (evaluation). The equilibrium may be sensitive hence not robust with respect to the way how the payoffs are evaluated. One may ask for the existence of an approximately equilibrium strategy for all repeated games with sufficiently patient discount factors (evaluation). This corresponds to the so-called *uniform equilibrium* we shall study.

Definition 3.3 *σ is a **uniform equilibrium** if for any sequence of patient evaluations $(\theta^k)_{k\geq 1}$:*
1) $\gamma_{\theta^k}(\sigma) \longrightarrow_{k\to+\infty} y$ for some y;
2) for any $\varepsilon > 0$, there exists some $K_0 \in \mathbb{N}$ such that for all $k \geq K_0$, σ is an ε-Nash equilibrium in the repeated game Γ_{θ^k}.

*σ is a **subgame perfect uniform equilibrium** if it is a uniform equilibrium in all subgames.*

Theorem 3.4 ***(Uniform Perfect Folk Theorem)*** *The set of perfect uniform equilibrium payoff set coincides with the individual rational set. To be precise: let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations and let x be any payoff vector in $\mathcal{F}$. Then for* *any small $\varepsilon > 0$, there is some behavior strategy $\sigma \in \Sigma$ such that:*

1) $\gamma_{\theta^k}(\sigma) \longrightarrow_{k\to+\infty} y$ for some $y \in \mathcal{F}$ with $||y - x||_\infty \le \varepsilon$;

2) there exists some $K_0 \in \mathbb{N}$ satisfying: for all $k \ge K_0$, σ is an subgame perfect ε-Nash equilibrium in the repeated game Γ_{θ^k}, i.e., $\forall T \ge 1$, $\forall h_T \in \mathcal{H}_T$, and $\forall \tau^i \in \Sigma^i$, $\forall i \in I$,

$$\gamma^i_{\theta^k[T]}(\sigma^i|_{h_T}, \sigma^{-i}|_{h_T}) \ge \gamma^i_{\theta^k[T]}(\tau^i|_{h_T}, \sigma^{-i}|_{h_T}) - \varepsilon.$$

Analogously, we obtain the following equivalent form of the above result.

Theorem 3.5 ***(Uniform Perfect Folk Theorem restated)*** *Let x be any payoff vector in $\mathcal{F}$. Then for any small $\varepsilon > 0$, there is some $\delta_0 \in (0, 1)$ and some behavior strategy $\sigma \in \Sigma$ such that: for all $\Lambda = (\delta_t)_{t\ge1}$ satisfying $\delta_t \in [\delta_0, 1), \forall t \ge 1$,*

1) $||\gamma_\Lambda(\sigma) - x||_\infty \le \varepsilon$;

2) σ is an subgame perfect ε-Nash equilibrium in the repeated game Γ_Λ, i.e., $\forall T \ge 1$, $\forall h_T \in \mathcal{H}_T$, and $\forall \tau^i \in \Sigma^i$, $\forall i \in I$,

$$\gamma^i_{\Lambda[T]}(\sigma^i|_{h_T}, \sigma^{-i}|_{h_T}) \ge \gamma^i_{\Lambda[T]}(\tau^i|_{h_T}, \sigma^{-i}|_{h_T}) - \varepsilon.$$

4 Preliminaries

We first give an equivalent characterization for the patient evaluations.

Lemma 4.1 $(\theta^k)_{k\ge1}$ *being a sequence of patient evaluations if and only if: for any fixed T_1 and T_2 with $0 \le T_1 \le T_2$,*

$$\forall T \in \mathbb{N}^*, \ \lim_{k\to\infty} \frac{\sum_{t=T}^{T+T_1} \theta^k_t}{\sum_{t=T}^{T+T_2} \theta^k_t} = \frac{T_1 + 1}{T_2 + 1}, \quad \textit{and the convergence is uniform in } T.$$

Proof. One direction is trivial if we take $T_1 = 0$ and $T_2 = 1$. Now we assume $(\theta^k)_{k\ge1}$ to be patient and let $T_2 \ge T_1 \ge 0$. For a fixed $\varepsilon > 0$, let $k \ge 1$ such that $1 \ge \theta^k_{t+1}/\theta^k_t \ge 1-\varepsilon$ for all $t \in \mathbb{N}^*$. By iteration, this implies that $1 \ge \theta^k_{t+s}/\theta^k_t \ge (1-\varepsilon)^s$ for all $s, t \ge 1$. For any $T \ge 1$, we bound the fraction $\frac{\sum_{t=T}^{T+T_1} \theta^k_t}{\sum_{t=T}^{T+T_2} \theta^k_t}$ as follows: first, we divide θ^k_T on both numerator and denominator; then, we replace each term $\theta^k_{T+s}/\theta^k_T$ by either $(1-\varepsilon)^s$ or 1. This yields:

$$\frac{\varepsilon(T_1+1)}{1-(1-\varepsilon)^{T_2+1}} = \frac{T_1+1}{\sum_{t=0}^{T_2}(1-\varepsilon)^t} \geq \frac{\sum_{t=T}^{T+T_1}\theta^k_t}{\sum_{t=T}^{T+T_2}\theta^k_t} \geq \frac{\sum_{t=0}^{T_1}(1-\varepsilon)^t}{T_2+1} = \frac{1-(1-\varepsilon)^{T_1+1}}{\varepsilon(T_2+1)}.$$

Since both $\frac{1-(1-\varepsilon)^{T_1+1}}{\varepsilon(T_2+1)}$ and $\frac{\varepsilon(T_1+1)}{1-(1-\varepsilon)^{T_2+1}}$ tend to $\frac{T_1+1}{T_2+1}$ as ε tends to zero, this proves that: $\forall T \in \mathbb{N}^*,\ \lim_{k\to\infty} \frac{\sum_{t=T}^{T+T_1}\theta^k_t}{\sum_{t=T}^{T+T_2}\theta^k_t} = \frac{T_1+1}{T_2+1}$. Now the choice of k is independent of T, so the convergence is uniform in T. ■

Not the same as the stationary discounting, the continuation distribution at stage $T \geq 1$ of a fixed evaluation is in general different for different T. This complicates the analysis, nevertheless, we have the following result which implies that the continuation sequence of evaluations of a sequence of patient evaluations is also patient.

Lemma 4.2 *Let $(\theta^k)_{k\geq 1}$ being a sequence of patient evaluations. Then it satisfies:*

1. $\forall T \in \mathbb{N}^*,\ \theta^k_1[T] \xrightarrow{k\to\infty} 0,$ *and the convergence is uniform in T;*
2. $\forall T, t \in \mathbb{N}^*,\ \frac{\theta^k_t[T]}{\theta^k_{t+1}[T]} = \frac{\theta^k_{T-1+t}[T]}{\theta^k_{T+t}[T]} \xrightarrow{k\to\infty} 1,$ *and the convergence is uniform in both T and t.*

Proof. Let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations, and we fix $T, t \geq 1$.

1. Following Lemma 4.1, for a fixed $\varepsilon > 0$, we let $T_1 = 0$ and $T_2 = \lfloor 2/\varepsilon \rfloor - 1$, then there is some $k \geq 1$ such that:

$$\forall T \geq 1,\ \left|\frac{\theta^k_T}{\sum_{t=T}^{T+T_2}\theta^k_t} - \frac{1}{T_2+1}\right| \leq \frac{\varepsilon}{2}, \quad \text{thus} \quad \theta^k_1[T] = \frac{\theta^k_T}{\sum_{t\geq T}\theta^k_t} \leq \frac{\theta^k_T}{\sum_{t=T}^{T+T_2}\theta^k_t} \leq \varepsilon.$$

2. The second result follows directly the definition. By definition, $\theta^k_t[T] = \frac{\sum_{s\geq t}\theta^k_{T-1+t}}{\sum_{s\geq t}\theta^k_s}$. Since $(\theta^k)_{k\geq 1}$ is patient,

$$\frac{\theta^k_t[T]}{\theta^k_{t+1}[T]} = \frac{\theta^k_{T-1+t}[T]}{\theta^k_{T+t}[T]} \xrightarrow{k\to\infty} 1, \text{ and the convergence is uniform in both } T \text{ and } t.$$

This proves the lemma. ■

Being analogue to Lemma 2 in [11], the following result plays an important role in the proof on the dispensation of public randomization.

Proposition 4.3 *Let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations, and let x be a payoff vector in $\mathcal{F}$. For any $\eta > 0$, there exists some $K(\eta)$ such that for any $k \geq K(\eta)$, there is a sequence of pure action profiles $h_\infty = (a_t)_{t\geq 1}$ satisfying*

i) $\gamma_{\theta^k}(h_\infty) = x$;

ii) after any stage $T \geq 1$, $||\gamma_{\theta^k[T+1]}(h_\infty^T) - x||_\infty \leq \eta$,

where h_∞ is decomposed as $(a_1, ..., a_T, h_\infty^T)$ and h_∞^T is the continuation history of some finite action sequence $(a_1, ..., a_T) \in A^T$.

Moreover, such integer $K(\eta)$ can be chosen independent of x.

Proof of Proposition 4.2: see the Appendix 7.2.

Notation. $B_\varepsilon(z) = \left\{y \in \mathbb{R}^I \middle| \|y - z\|_\infty \leq \varepsilon\right\}$, $M = \max\{|g^i(a)| : i \in I, a \in A\}$.

The proof of Proposition 4.2 involves the following technical result.

Lemma 4.4 *Let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations, and let x be a convex combination of a set of vectors $\{x_\ell : 1 \leq \ell \leq L\}$ in $\mathbb{R}^I$ with $L \leq I + 1$ satisfying: for any ℓ, the set of vectors $\{x_{\ell'} - x_\ell | \ell' \neq \ell\}$ are linear independent. Consider any small $\eta > 0$ with $B_\eta(x) \subseteq co\{x_\ell\}$ and let $\eta' = \eta/5$. Then there is some $K(\eta)$ such that: for any $k \geq K(\eta)$ the following is satisfied: for any fixed set of vectors $\{x_{\ell,m} : 1 \leq \ell \leq L, m \geq 1\}$ with each $x_{\ell,m} \in B_{\eta'}(x_\ell)$, one can find a sequence of vectors $\left(x'_{\varphi(m)}\right)_{m\geq 1}$ with each $x'_{\varphi(m)}$ taking value in $\{x_{\ell,\varphi(m)} : 1 \leq \ell \leq L\}$ such that $\sum_{m=1}^\infty \theta_m^k x'_{\varphi(m)} = x$.*

Note that our previous result is a perturbed, hence generalized, version of Proposition 4 in [26] or of Lemma 2 in [11]. In the stationary discounting case, it is sufficient to get the result for $x_{\ell,m} = x_\ell$. We see later how this perturbation is used.

Proof of Lemma 4.4: see the Appendix 7.1.

5 Proofs of the main results

5.1 Asymptotic analysis: equilibrium construction and optimality check

This section is devoted to the proof of Theorem 3.1. In **Part A** we specify some parameters to be used. **Part B** is the definition of the equilibrium strategy. In **Part C**, we check the optimality.

Let x be a payoff vector in the interior of $\mathcal{F}$. Since $\mathcal{F}$ has a fully dimensional nonempty interior and is convex, there exist $w \in int\mathcal{F}$ and a sufficiently small $\varepsilon > 0$ satisfying $B_\varepsilon(w) \subseteq \mathcal{F}$ and $x^i - w^i \geq 2\varepsilon$ for all $i \in I$. For each player $i \in I$ and for a fixed $z_i = (z_i^j)_{j\in I} \in \mathbb{R}^I$, we define the payoff vector $w[z_i] = (w^j[z_i])_{j\in I}$ to be: $w^j[z_i] = w^j + \left(\varepsilon - z_i^j\right)\mathbb{1}_{j\neq i}$, $\forall j \in I$. As long as $||z_i||_\infty \leq \varepsilon/2$ for all $i \in I$, we obtain $w[z_i] \in \mathcal{F}$ for all $i \in I$.

We fix below $(\theta^k)_{k\geq 1}$ a sequence of patient evaluations.

Part A. Specification of the parameters ε, η, k, L.

Denote $\kappa = \min\{w^i - v^i | i \in I\} > 0$. For the given $\varepsilon > 0$, we choose a positive number $\eta < \min\{\frac{\varepsilon}{8M+1}, \frac{1}{2M}, \frac{\kappa}{8}, \frac{\kappa\varepsilon}{48M+16\varepsilon+4\kappa}\}$. Now we fix the parameter $K(\eta)$ satisfying Proposition 4.3. Since $(\theta^k)_{k\geq 1}$ is patient, we can choose $k \geq K(\eta)$ sufficiently large such that (by Lemma 4.2)

$$\theta_1^k[T] \leq \eta^2, \text{ for all } T \geq 1.$$

For fixed η and k, we define for each $T \geq 1$ the positive integer $L(\eta, k, T)$ to satisfy

$$\frac{3\eta}{\kappa} \leq \sum_{t=1}^{L(k,\eta,T)+1} \theta_t^k[T] < \frac{4\eta}{\kappa}. \tag{5.1}$$

This is possible since each θ_t^k is decreasing in t, thus

$$\theta_t^k[T] \leq \theta_1^k[T] \leq \eta^2 < \frac{\eta}{\kappa}, \ \forall t \geq 1,$$

where we used the fact that $\eta < \frac{1}{2M} \leq \frac{1}{\kappa}$.

Part B. Construction of the equilibrium strategy.

The equilibrium strategy is constructed in a similar way as Fudenberg and Maskin (1986) which consists of three phases: the *main path* Phase I, the *punishing* Phase II and the *rewarding* Phase III. On the main path Phase I, players play cycles of pure actions that lead to the payoff vector x, and continue with the cyclical actions as long as there is no deviation. Once a deviation by some player i has been detected at some stage T, the play moves to the Phase II(i) in which players start to punish i for $L(\eta, k, T)$ stages. If Phase II(i) ends with no deviation, the play moves to the Phase III(i) rewarding $-i$ (all players other than i), which is defined by cycles of pure actions leading to a payoff vector $w[z_i]$ with everyone strictly higher than w^j except for player i, attaining only w^i. Any deviation by a player j within the punishing Phase II(i) or the rewarding Phase III(i) will again lead to a punishing Phase II(j).

As in Fudenberg and Maskin (1986), the introduction of the rewarding Phase III after the punishing Phase II with a bounded length is to obtain the *subgame perfection* of the equilibrium: other players have the sufficient incentive to commit punishing the deviated player. The careful specification of the length $L(\eta, k, T)$ in **Part A** is such that: on one hand it is large enough so any deviation from the main path is deterred; on the other hand, it is not too long so the incentive of rewards in Phase III is still effective.

Let the parameters η, k and the integer function $L(\eta, k, T)$ be specified as in **Part A**. For any stage number T and for each $i \in I$, define a vector mapping $z_i(\cdot)$ from $A^{L(\eta,k,T)}$ to $[-\varepsilon/2, \varepsilon/2]^I$:

$$\forall j \in I,\ z_i^j(a_1, ..., a_{L(\eta,k,T)}) = \Diamond \times p_i^j(a_1, ..., a_{L(\eta,k,T)}),$$

where $\Diamond := \frac{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}{1-\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}$ and $p_i^j(a_1, ..., a_{L(\eta,k,T)}) = \frac{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T] g^j(a_t)}{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}$ is the $\theta^k[T]$-evaluated average payoff for player j during the $L(\eta, k, T)$ periods over which the actions profiles $(a_1, ..., a_{L(\eta,k,T)})$ were played.

We first show that

Lemma 5.1 $\left|z_i^j(a_1, ..., a_{L(\eta,k,T)})\right| \leq \varepsilon/2, \ \ \forall j \in I.$

Proof. By construction, $z_i^j(a_1, ..., a_{L(\eta,k,T)}) = \frac{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}{1-\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]} \times p_i^j(a_1, ..., a_{L(\eta,k,T)})$ and p_i^j is a normalized average payoff, it is sufficient for us to prove that $\frac{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}{1-\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]} \leq \frac{\varepsilon}{2M}$. Indeed, from (5.1) one has $\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T] \leq \frac{4\eta}{\kappa}$. Finally, the result is obtained since we have chosen η small so both $1 - \frac{4\eta}{\kappa} > 1/2$ (as $\eta < \frac{\kappa}{8}$) and $\frac{4\eta}{\kappa} < \frac{\varepsilon}{4M}$ (as $\eta < \frac{\kappa\varepsilon}{16M}$) are satisfied. ∎

For a fixed vector $(a_1, ..., a_{L(\eta,k,T)})$, we write z_i^j shortly for $z_i^j(a_1, ..., a_{L(\eta,k,T)})$ and denote $z_i = (z_i^j)_{j\in I}$. Let us write also $w[z_0] = x$. Next we want to apply Proposition 4.3 to the vectors $w[z_0], w[z_1], ..., w[z_I]$ one by one. Notice that each $w[z_i]$ may correspond to different sequence of evaluations $(\theta^k[T_i])$ for different T_i and this may require different parameter $K(\eta)$. However, Lemma 4.2 ensures us that such $K(\eta)$ can be taken independent of $w[z_i]$. That is, for some $k \geq K(\eta)$, there are $(I + 1)$ sequences of pure actions $\mathbf{x}[0], \mathbf{x}[1], ..., \mathbf{x}[I]$ such that $\forall i = 0, 1, ..., I$,

1). $\gamma_{\theta^k}(\mathbf{x}[i]) = w[z_i]$;

2). $\left\|\gamma_{\theta^k[t]}(\mathbf{x}^t[i]) - w[z_i]\right\|_\infty \leq \eta$, $\forall t \geq 1$, where $\mathbf{x}^t[i]$ is the continuation history of $\mathbf{x}[i]$ after some $(a_1, ..., a_t) \in A^t$, i.e. $\mathbf{x}[i] = (a_1, ..., a_t, \mathbf{x}^t[i])$.

We define now some behavior strategy σ whose θ^k-evaluated payoff is x. Moreover, we prove that σ is a subgame perfect Nash equilibrium of the repeated game Γ_{θ^k}.

σ is defined on phases:

- Phase I (*main path*). σ starts from Phase I where players follow the history $\mathbf{x}[0]$ as long as there is no deviation. Suppose that there is some player, say i, who has deviated from this phase at some stage T, then the play moves to Phase II$\big(i; T; L(\eta, k, T)\big)$;

- Phase II$\big(i; T; L(\eta, k, T)\big)$ (*punishing*). On this phase, all players other than i punish player i over $L(\eta, k, T)$ periods, during which the minmax (mixed) action $\hat{s}^i$ is played i.i.d.

– If some player[11] $j \neq i$ deviated from $\hat{s}^i$ at some stage T' during the $L(\eta, k, T)$ periods, by which we mean that some of his pure action outside the support of $\hat{s}^i$ has been detected, then the play moves to Phase $\text{II}\big(j; T'; L(\eta, k, T')\big)$, punishing player j over $L(\eta, k, T')$ periods.

– Otherwise (no deviation during this phase), the play shifts to Phase $\text{III}(i; a_1, ..., a_{L(\eta,k,T)})$, where $(a_1, ..., a_{L(\eta,k,T)}) \in A^{L(\eta,k,T)}$ is the sequence of realized pure action profiles during Phase $\text{II}\big(i; T; L(\eta, k, T)\big)$.

- Phase $\text{III}(i; a_1, ..., a_{L(\eta,k,T)})$ (*rewarding*). On this phase, players follow the history $\mathbf{x}[i]$ $\big($which depends on $(a_1, ..., a_{L(\eta,k,T)})$ through the vector $z_i^j(a_1, ..., a_{L(\eta,k,T)})\big)$. Suppose that there is some player j (could be i or not) who has deviated from this phase at some stage T', then the play moves to Phase $\text{II}\big(j, T; L(\eta, k, T')\big)$.

Part C. Checking the optimality: no profitable deviation exists.

Now let us verify that no profitable deviation from σ exists. It is sufficient for us to consider the one-shot deviation[12] (cf. Theorem 4.2 in Fudenberg and Tirole [10]).

1. *The main path.* Suppose that the deviation occurs at some stage T during Phase I by player i. A sufficient condition for the deviation to be non-profitable is:

$$\theta_T^k M + \sum_{t=1}^{L(\eta,k,T)} \theta_{T+t}^k v^i + \Big(\sum_{t > T + L(\eta,k,T)} \theta_t^k\Big)(w^i + \eta) \leq \Big(\sum_{t \geq T} \theta_t^k\Big)(x^i - \eta),$$

   which is equivalent to $\big($by dividing $(\sum_{t \geq T} \theta_t^k)$ on both sides of the above inequality$\big)$

$$\theta_1^k[T] M + \sum_{t=2}^{L(\eta,k,T)+1} \theta_t^k[T] v^i + \Big(\sum_{t > L(\eta,k,T)+1} \theta_t^k[T]\Big)(w^i + \eta) \leq x^i - \eta \qquad (5.2)$$

[11] Note that there is no need for σ to specify player i's actions during this phase.

[12] As in the stationary discounting case (cf. Proposition 2.21 in Mailath and Samuelson [17]), the principle of one-shot deviation also holds true under general discounting. In fact, the key step is the property of "*continuity at infinity*", i.e. the action in the far future has little impact on the total payoff: for any evaluation $(\theta_t)_{t \geq 1}$ and any $\varepsilon > 0$, there exists some $T > 0$ such that $\sum_{t > T} \theta_t < \varepsilon$.

The LHS of Inequality (5.2) is bounded from above by

$$2M\Big(\sum_{t=1}^{L(\eta,k,T)+1}\theta_t^k[T]\Big)+w^i+\eta\le\big(\frac{8M}{\kappa}+1\big)\eta+w^i,$$

where we have used the fact that $\sum_{t=1}^{L(\eta,k,T)+1}\theta_t^k[T]\le\frac{4\eta}{\kappa}$ according to (5.1). Next, since we have chosen $\eta<\frac{\kappa\varepsilon}{48M+16\varepsilon+4\kappa}<\frac{\kappa\varepsilon}{8M+\kappa}$, it implies that $\big(\frac{8M}{\kappa}+1\big)\eta<\varepsilon$. So we obtain that the LHS of (5.2) is bounded by $\varepsilon+w^i$, which is smaller than the RHS of (5.2) since $x^i-w^i\ge 2\varepsilon$ and $\eta<\frac{\varepsilon}{8M+1}<\varepsilon$. This proves (5.2), thus any one-shot deviation within this phase is not profitable.

2. *The rewarding phase.* Suppose that we are in some Phase III$(i;a_1,...,a_L)$ with respect to some player i, where $(a_1,...a_L)$ are the realized pure actions during the preceding punishing over L periods. The equilibrium strategy assigns players to follow the sequence of pure action profiles $\mathbf{x}[i]$ leading to a θ^k-evaluated payoff η close to $w[z_i]$.

   Consider now a deviation from $\mathbf{x}[i]$ made by some player j (could be i or not) happens on this phase, say at some stage T. Then a sufficient condition for this deviation to be non-profitable is:

$$\theta_T^kM+\sum_{t=1}^{L(\eta,k,T)}\theta_{T+t}^kv^j+\Big(\sum_{t>T+L(\eta,k,T)}\theta_t^k\Big)(w^j+\eta)\le\Big(\sum_{t\ge T}\theta_t^k\Big)(w^j[z_i]-\eta),$$

   which is equivalent to

$$\theta_1^k[T]M+\sum_{t=2}^{L(\eta,k,T)+1}\theta_t^k[T]v^j+\Big(\sum_{t>L(\eta,k,T)+1}\theta_t^k[T]\Big)(w^j+\eta)\le w^j[z_i]-\eta.$$

   By definition, $w^j[z_i]=w^j+(\varepsilon-z_i^j)\mathbb{1}_{j\neq i}\ge w^j$ for all $j\in I$, thus it is sufficient to prove:

$$\theta_1^k[T]M+\sum_{t=2}^{L(\eta,k,T)+1}\theta_t^k[T]v^j+\Big(\sum_{t>L(\eta,k,T)+1}\theta_t^k[T]\Big)(w^j+\eta)\le w^j-\eta. \tag{5.3}$$

The LHS of the above inequality (5.3) is equal to:

$$\theta_1^k[T](M - v^j) + \Big(\sum_{t=1}^{L(\eta,k,T)+1} \theta_t^k[T]\Big)(v^j - w^j) + w^j + \Big(\sum_{t>L(\eta,k,T)+1} \theta_t^k[T]\Big)\eta$$
$$\leq 2M\theta_1^k[T] - \Big(\sum_{t=1}^{L(\eta,k,T)+1} \theta_t^k[T]\Big)\kappa + w^j + \eta.$$

To obtain (5.3), it is sufficient for us to prove that

$$2M\theta_1^k[T] + 2\eta \leq \Big(\sum_{t=1}^{L(\eta,k,T)+1} \theta_t^k[T]\Big)\kappa. \tag{5.4}$$

According to (5.1), we have on the RHS of (5.4): $\Big(\sum_{t=1}^{L(\eta,k,T)+1} \theta_t^k[T]\Big)\kappa \geq 3\eta$. On the LHS of (5.4): k is chosen large with $\theta_1^k[T] \leq \eta^2$ and moreover $\eta < \frac{1}{2M}$, so we obtain $2M\theta_1^k[T] + 2\eta \leq 2M\eta^2 < \eta$. This proves (5.4), thus (5.3), so any one-shot deviation within this phase is not profitable.

3. *The punishing phase.* Suppose that the play is in some Phase $\mathrm{II}\big(i; T; L(\eta, k, T)\big)$, which implement the punishment over player i, started from stage $T+1$ and will last for $L(\eta, k, T)$ periods. Consider now a one-shot deviation by some player $j \neq i$ from playing the minmax strategy $\hat{s}^i$ (which could be mixed).

   1). We first argue that by our construction of the contingent payoff function $z_i^j(a_1, ..., a_{L(\eta,k,T)})$, any one-shot deviation within the support of $\hat{s}^i$ (not detectable) is not profitable. Indeed, according Proposition 4.2 and also the construction of the equilibrium path on the rewarding phase, player j's realized payoff over Phase $\mathrm{II}\big(i; T; L(\eta, k, T)\big)$ is $\Big($recall that $z_i^j(a_1, ..., a_{L(\eta,k,T)}) = \frac{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}{1-\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]} p_i^j$, where $p_i^j := p_i^j(a_1, ..., a_{L(\eta,k,T)})\Big)$:

$$\Big(\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]\Big)p_i^j + \Big(1 - \sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]\Big)\left[w^j + \varepsilon - \frac{\sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]}{1 - \sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]} p_i^j\right]$$
$$= \Big(1 - \sum_{t=1}^{L(\eta,k,T)} \theta_t^k[T]\Big)\Big(w^j + \varepsilon\Big),$$

which is independent of the sequence of realized pure action profiles $(a_1, ..., a_{L(\eta,k,T)})$. In particular, player j's expected payoff is independent of his strategy as long as at each stage his realized action is within the support of the minmax strategy $\hat{s}^i$.

2). We then consider player j's deviation outside the support of $\hat{s}^i$, say at some stage $T' \in \{T+1, T+L(\eta,k,T)\}$ during Phase II$\big(i; T; L(\eta,k,T)\big)$. Let us denote $S = T + L(\eta,k,T) - T'$, which is the number of periods from stage $T'+1$ on and till the end of this phase. The play shall shift to Phase II$\big(j; T'; L(\eta,k,T')\big)$ periods after the deviation by j. A sufficient condition for this one-shot deviation to be non-profitable is:

$$\begin{aligned} &\theta^k_{T'} M + \sum_{t=1}^{L(\eta,k,T')} \theta^k_{T'+t} v^j + \Big(\sum_{t > T + L(\eta,k,T')} \theta^k_t \Big)(w^j + \eta) \\ &\leq \mathbb{E}_\sigma \left[\sum_{t=0}^{S} \theta^k_{T'+t} u^j(\tilde{a}_{T'+t}) + \Big(\sum_{t > T'+S} \theta^k_t \Big) \Big(w^j + \varepsilon - \tilde{z}^j_i - \eta \Big) \Big| \mathcal{H}_{T'} \right], \end{aligned} \tag{5.5}$$

where $\mathcal{H}_{T'}$ is the information set at (the start of) stage T' and $\mathbb{E}_\sigma[\cdot|\mathcal{H}_{T'}]$ is the conditional expectation of the remaining play, $\tilde{a}_{T'+t}$ is the action profile (random variable) to be played in stage $T'+t$, and $\tilde{z}^j_i$ refers to the random variable $z^j_i[a_{T+1}, ..., a_{T'-1}, \tilde{a}_{T'}, ..., \tilde{a}_{T'+S}]$, defined as a function of the sequence of realized pure actions within Phase II$\big(i; T; , L(\eta,k,T)\big)$.

Below we prove (5.5) in two cases[13].

<u>Case One: $L(\eta,k,T') \geq S$.</u>

[13] In the stationary discounting case, $\theta_t[T] = \theta_t[T']$ for all T, T', t, thus we could have chosen $L(\eta,k,T)$ independent of T. As a consequence, we always have $S \leq L(\eta,k,T) = L(\eta,k,T')$, and there is no need for us to discuss by cases. One is referred to Proposition 3.8.1 in Mailath and Samuelson [17] for the proof of stationary discounting case.

We observe that

$$\text{RHS of (5.5)} \geq \Big(\sum_{t=0}^{L(\eta,k,T')} \theta^k_{T'+t}\Big)(-M) + \Big(\sum_{t>L(\eta,k,T')} \theta^k_{T'+t}\Big)(w^j + \varepsilon/2 - \eta)$$

$$\geq \Big(\sum_{t=0}^{L(\eta,k,T')} \theta^k_{T'+t}\Big)(-M - w^j - \varepsilon/2 + \eta) + \Big(\sum_{t\geq T'} \theta^k_t\Big)(w^j + \varepsilon/2 - \eta)$$

where we have used the bound "$u^j(a) \geq -M,\ \forall a \in A$" and the fact that "$|z^j_i| \leq \varepsilon/2,\ \forall j$" for any realization of the pure actions (Lemma 5.1). On the other hand,

$$\text{LHS of (5.5)} \leq \Big(\sum_{t=0}^{L(\eta,k,T')} \theta^k_{T'+t}\Big)M + \Big(\sum_{t>L(\eta,k,T')} \theta^k_{T'+t}\Big)(w^j + \eta)$$

$$\leq \Big(\sum_{t=0}^{L(\eta,k,T')} \theta^k_{T'+t}\Big)(M - w^j - \eta) + \Big(\sum_{t\geq T'} \theta^k_t\Big)(w^j + \eta)$$

By dividing "$(\sum_{t\geq T'} \theta^k_t)$" on both sides, the above computations tell us that (5.5) is implied by:

$$\Big(\sum_{t=0}^{L(\eta,k,T')} \theta^k_{T'+t}\Big)(2M + \varepsilon/2 - 2\eta) \leq \Big(\sum_{t\geq T'} \theta^k_t\Big)(\varepsilon/2 - 2\eta)$$

$$\Longleftarrow \Big(\sum_{t=1}^{L(\eta,k,T')+1} \theta^k_t[T']\Big)(2M + \varepsilon/2) + 2\eta \leq \varepsilon/2.$$

Since we have chosen the integer $L(\eta, k, T')$ such that $\Big(\sum_{t=1}^{L(\eta,k,T')+1} \theta^k_t[T']\Big) \leq \frac{4\eta}{\kappa}$ $\big($following (5.1)$\big)$ and we have set $\eta < \frac{\kappa\varepsilon}{48M+16\varepsilon+4\kappa} < \frac{\kappa\varepsilon}{16M+4\varepsilon+4\kappa}$, an easy computation proves the above inequality hence (5.5).

Case Two: $L(\eta, k, T') < S$.

The proof is more involved in this case. We first make the following computation:

RHS - LHS of (5.5)

$$\geq \Big(\sum_{t=0}^{L(\eta,k,T')} \theta^k_{T'+t}\Big)(-2M) + \Big(\sum_{t>L(\eta,k,T')}^{S} \theta^k_{T'+t}\Big)(-2M - \eta) + \Big(\sum_{t>S} \theta^k_{T'+t}\Big)(\varepsilon/2 - 2\eta),$$

and we divide "$(\sum_{t \geq T'} \theta_t^k)$" on both sides of the above inequality to obtain

$$\frac{\text{RHS - LHS of (5.5)}}{\sum_{t \geq T'} \theta_t^k} \geq \alpha(1, L)[-2M] + \alpha(L, S)[-2M - \eta] + \alpha(S, \infty)[\varepsilon/2 - 2\eta], \tag{5.6}$$

where $\alpha(1, L) := \left(\sum_{t=1}^{L(\eta,k,T')+1} \theta_t^k[T']\right)$, $\alpha(L, S) := \left(\sum_{t>L(\eta,k,T')+1}^{S+1} \theta_t^k[T']\right)$ and $\alpha(S, \infty) := \left(\sum_{t>S+1} \theta_s^k[T']\right)$. To prove (5.5), it is sufficient for us to show that

$$\alpha(1, L)[-2M] + \alpha(L, S)[-2M - \eta] + \alpha(S, \infty)[\varepsilon/2 - 2\eta] \geq 0.$$

Since $\alpha(1, L) + \alpha(L, S) + \alpha(S, \infty) = 1$, the above inequality is equivalent to

$$\alpha(1, L)[-2M - \varepsilon/2 + 2\eta] + \alpha(L, S)[-2M - \eta - \varepsilon/2 + 2\eta] + [\varepsilon/2 - 2\eta] \geq 0,$$

which is implied by

$$\begin{aligned} &\alpha(1, L)[-2M - \varepsilon/2] + \alpha(L, S)[-2M - \varepsilon/2] + [\varepsilon/2 - 2\eta] \geq 0 \\ \iff &\varepsilon/2 \geq 2\eta + \alpha(1, L)[2M + \varepsilon] + \alpha(L, S)[2M + \varepsilon/2] \end{aligned} \tag{5.7}$$

Since we have taken η small compared to ε, $\alpha(1, L) = \left(\sum_{t=1}^{L(\eta,k,T')+1} \theta_t^k[T']\right)$ is bounded by $4\eta/\kappa$, it is sufficient for us to prove that $\alpha(L, S) = \left(\sum_{t>L(\eta,k,T')+1}^{S+1} \theta_t^k[T']\right)$ is also small (bounded by η for example).

Lemma 5.2 $\alpha(L, S) = \left(\sum_{t>L(\eta,k,T')+1}^{S+1} \theta_t^k[T']\right) \leq \frac{8\eta}{\kappa}$.

Proof. During the proof of the lemma, we denote $L := L(\eta, k, T)$ and $L' := L(\eta, k, T')$. By definition, we have (by the facts that $T' + S = T + L$ and $T < T' + L' + 1$):

$$\alpha(L, S) = \left(\sum_{t>L(\eta,k,T')+1}^{S+1} \theta_t^k[T']\right) = \frac{\sum_{t=T'+L'+1}^{T'+S} \theta_t^k}{\sum_{t \geq T'} \theta_t^k} \leq \frac{\sum_{t=T}^{T+L} \theta_t^k}{\sum_{t \geq T} \theta_t^k} \cdot \frac{\sum_{t \geq T} \theta_t^k}{\sum_{t \geq T'} \theta_t^k}.$$

We look at the two terms on the RHS of the above inequality separately:

- first, we have $\frac{\sum_{t=T}^{T+L} \theta_t^k}{\sum_{t \geq T} \theta_t^k} = \sum_{t=1}^{L+1} \theta_t^k[T]$, which is smaller than $4\eta/\kappa$, by the definition of $L := L(\eta, k, T)$ in (5.1);

- second, we write

$$\frac{\sum_{t\geq T}\theta_t^k}{\sum_{t\geq T'}\theta_t^k} = \frac{\sum_{t\geq T}\theta_t^k}{\sum_{t\geq T}\theta_t^k - \sum_{T\leq t<T'}\theta_t^k} \leq \frac{\sum_{t\geq T}\theta_t^k}{\sum_{t\geq T}\theta_t^k - \sum_{t=T}^{T+L}\theta_t^k} = \frac{1}{1-\sum_{t=1}^{L+1}\theta_t^k[T]}.$$

The above computations imply that (using the bound "$\eta < \frac{\kappa}{8}$")

$$\alpha(L,S) \leq \frac{\sum_{t=1}^{L+1}\theta_t^k[T]}{1-\sum_{t=1}^{L+1}\theta_t^k[T]} \leq \frac{4\eta}{\kappa-4\eta} < \frac{8\eta}{\kappa}.$$

■

Now we substitute the bounds "$\alpha(1,L) \leq \frac{4\eta}{\kappa}$" and "$\alpha(L,S) \leq \frac{8\eta}{\kappa}$" back into (5.7) to obtain that it is implied by:

$$\varepsilon/2 \geq 2\eta + \frac{4\eta}{\kappa}(2M+\varepsilon) + \frac{8\eta}{\kappa}(2M+\varepsilon/2).$$

It is easy to verify that by the bound "$\eta < \frac{\kappa\varepsilon}{48M+16\varepsilon+4\kappa}$" as was specified, the above inequality is satisfied. Then (5.7) is proved and (5.6), hence (5.5), is proved.

This finishes the proof that under both cases, there is no profitable one-shot deviation within any punishing Phase II$(i;T,L(\eta,k,T))$.

We have thus verified that there is no profitable one-shot deviation within any phase, so σ is a subgame perfect Nash equilibrium of the repeated game Γ_{θ^k} fir a sufficiently large k. Moreover, the expected payoff under σ is η-close to the vector $x \in \Delta$. As $\eta > 0$ can be taken arbitrarily small, it means that as k tends to infinity, E_{θ^k} converges to Δ with respect to the Hausdoff distance.

This finishes the proof of Theorem 3.1.

5.2 Uniform analysis: equilibrium construction and optimality check

This section is devoted to the proof of Theorem 3.4.

The proof for the uniform perfect folk theorem is much simpler than the asymptotic one. The construction of the uniform equilibrium is as in Aumann and Shapley[2] and as in Rubinstein [23] for the undiscounted criterion.

Let x be a payoff vector in the individual rational payoff set $\mathcal{F}$, i.e. $x^i \geq v^i$, $\forall i \in I$. We fix $\varepsilon > 0$ and let $\eta < \varepsilon/3$ be a small positive number. $(\theta^k)_{k\geq 1}$ is a sequence of patient evaluations.

We first prove the following result which takes a very close form to Proposition 4.3. It is weaker than Proposition 4.3 in the sense that $\gamma_{\theta^k}(h_\infty)$ does not lead to x exactly, and it is stronger than Proposition 4.3 in the sense that such h_∞ is taken independent of k.

Proposition 5.3 *Let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations, and let x be a payoff vector in $\mathcal{F}$. For any $\eta > 0$, there is some $K(\eta)$ and a sequence of pure action profiles $h_\infty = (a_t)_{t\geq 1}$ such that: $\forall k \geq K(\eta)$,*
i) $||\gamma_{\theta^k}(h_\infty) - x||_\infty \leq \eta$;
ii) after any stage $T \geq 1$, $||\gamma_{\theta^k[T+1]}(h^T_\infty) - x||_\infty \leq \eta$,
where h_∞ is decomposed as $(a_1, ..., a_T, h^T_\infty)$ and h^T_∞ is the continuation history of some finite action sequence $(a_1, ..., a_T) \in A^T$.
Moreover, such integer $K(\eta)$ can be chosen independent of x.

Proof of Proposition 5.3: see the Appendix 7.3.

We let $K(\eta)$ be the specified number and h_∞ be the sequence of pure action profiles as in Proposition 5.3 that "leads" to x. First it is easy to observe that $\gamma_{\theta^k}(h_\infty) \xrightarrow{k\to\infty} y$ for some $y \in \mathbb{R}^I$.

Next, fix some integer number $\bar{R} > 2M/\eta$, and choose $K(\eta)$ even larger satisfying: for all $k \geq K(\eta)$ and $T \geq 1$,

$$\theta^k_1[T] \leq \frac{\eta}{2M(\bar{R}+1)} \ \big(\text{Lemma 4.2}\big) \ \text{ and } \ \left|\frac{\theta^k_1[T]}{\sum_{t=1}^{\bar{R}+1}\theta^k_t[T]} - \frac{1}{1+\bar{R}}\right| \leq \frac{\eta}{2M} \ \big(\text{Lemma 4.1}\big).$$

The equilibrium strategy profile σ is constructed as follows.

- On the *main path*, players follow h_∞ and continue with it as long as no deviation occurs;

- Once a deviation by player i from the main path has been detected, the other players $-i$ start punishing the deviated player i for $\bar{R}$ stages, and then return to the main path whatever happened during the punishing phase.

That is to say, any deviation from the main path is followed by a punishing phase of $\bar{R}$ stages and the punishment is only designed for deviation during the main path. Below we verify that σ is a subgame perfect uniform equilibrium of the repeated game. For this aim, we check that any deviation will lead to a profit no more than ε in any subgame of Γ_{θ^k} provided that k is sufficiently large.

We first look at any possible deviation during the main path. By the specification of $K(\eta)$ and punishing length $\bar{R}$, after deviation at any stage T by player i, the weighted average payoff on the stages $T+1, ..., T+\bar{R}$ for him is, for $k \geq K(\eta)$, at most

$$\frac{\theta_1^k[T]M + \sum_{t=2}^{\bar{R}+1} \theta_t^k[T]v^i}{\sum_{t=1}^{\bar{R}+1} \theta_t^k[T]} \leq \frac{M + \bar{R}v^i}{1+\bar{R}} + \eta \leq v^i + \eta,$$

so the profit is at most $(v^i + \eta) - (x^i - \eta) \leq 2\eta$.

Next, we consider deviation by player $j \neq i$ during the phase punishing player i (starting from stage $T+1$). As the punishing phase of $\bar{R}$ periods is uniformly bounded, the profit from this deviation is at most $2M\big(\bar{R}+1\big)\theta_1^k[T] \leq \eta$ for $k \geq K(\eta)$.

Consider now any deviation from σ^i that could be a combination of the above two types, the associated total profit will be at most $3\eta < \varepsilon$.

This proves that σ is a subgame perfect uniform equilibrium, i.e. 1) it leads to a payoff vector ε close to x; and 2) it is an ε-Nash equilibrium in any subgame of all repeated games Γ_{θ^k} with $k \geq K(\eta)$. Since x is any payoff vector in $\mathcal{F}$ and $\varepsilon > 0$ can be chosen arbitrarily small, this finishes our proof of Theorem 3.4.

6 Time-inconsistent preferences

In this section we discuss a different perspective of time-dependent discounting that allows for time inconsistency and includes the present-biased preference as a special case. Let $\delta_{T,t}$ be the discount rate for payoff in stage t to stage $t-1$ when the player is currently at stage T (assuming $t \geq T$). That is, as the game proceeds, players' time preference over payoffs in stage t to stage $t-1$ evolves. This behavioral perspective is conceptually different from the rational one.

For each $T \geq 1$, let $\theta(T) \in \Delta(\mathbb{N}^*)$ be the evaluation associated with the discount sequence $(\delta_{T,t})_{t\geq 1}$ when the game is currently at stage T. The time-consistent preference actually means $\theta(T) = \theta[T]$, while the time-inconsistent preference does not assume this and $\theta(T)$ can be arbitrary. One special case is that $\theta(T) = \theta(T') := \theta$ for different T, T'. This is the time preference discussed in Obara and Park [20]

Example 2 *For $\beta, \delta \in (0,1)$, the (β,δ)-preference (hyberbolic discounting) has a series of discount factors for future payoffs $(1, \beta\delta, \beta^2\delta, ..., \beta\delta^t, ...)$ for a subgame at any stage T, which can be characterized by the evaluation $\theta_1 = \frac{1}{\eta}$ and $\theta_t = \frac{\beta\delta^{t-1}}{\eta}$, $\forall t \geq 2$, where $\eta = 1 + \beta\sum_{t\geq 2}\delta^{t-1} = \frac{1-\delta+\beta\delta}{1-\delta}$. Consider now a choice between a) \$ 50 realizing now and \$ 100 realizing the next stage, and b) \$ 50 realizing in stage 90 stages from now and \$ 100 realizing 91 stages from now. Assuming $\delta\beta < 1/2 < \delta$, the decision maker with such a preference will choose \$ 50 today in a) and \$ 100 in b). Time inconsistency emerges when the play arrives at stage 90 since now she/he prefers the \$ 50 realizing now.*

Now we turn to folk theorems under time-inconsistent preferences. For this aim, we need first to talk about the subgame perfect Nash equilibrium concept in this context. With time-inconsistent preference, there are multi-selves acting at different stages. By multi-selves, we mean that upon making a decision at stage T, this "self" wants to optimize the future payoffs evaluated by $(\delta_{T,t})_{t\geq T}$. Following the literature[14], we adopt the *Strotz-Pollak* equilibrium concept for the subgame perfect Nash equilibrium. That is, each self believes that her/his future selves would stick

[14]Initially discussed by Strotz [25] for dynamic decision problems, formalized by Peleg and Yaari [22], and introduced and studied for repeated games by Chade et al. [7]. In Obara and Park [20], it is called the *agent subgame perfect Nash equilibrium.*

to the specified equilibrium strategy when she/he deviates in the current stage. The idea is that different selves are generally endowed with different objective functions, thus may not coordinate with her/him.

The formal definition is as follows.

Notation. Let us denote $\Lambda(T) := (\delta_{T,t})_{t\geq 1}$. Interchangably, we may use either $\gamma_{\Lambda(T)}(\sigma)$ or $\gamma_{\theta(T)}(\sigma)$ (*resp.* either $\Gamma_{\{\Lambda(T)\}}$ or $\Gamma_{\{\theta(T)\}}$, either $E_{\{\Lambda(T)\}}$ or $E_{\{\theta(T)\}}$) for the expected payoff (*resp.* repeated game, perfect equilibrium payoff set). $(a^i, \sigma^i)|_{h_T})$ refers to player i's strategy in the subgame $\Gamma(h_T)$ that takes action a^i in the first stage and follows σ^i thereafter.

Definition 6.1 *A strategy profile $\sigma \in \Sigma$ is a **subgame perfect Nash equilibrium** (resp. **subgame perfect ε-Nash equilibrium**, $\varepsilon > 0$) of the repeated game $\Gamma_{\{\theta(T)\}}$ if for all history $h_T \in \mathcal{H}_T, T \geq 1$, $\forall a^i \in A^i, \forall i \in I$,*

$$\gamma^i_{\theta(T)}(\sigma^i|_{h_T}, \sigma^{-i}|_{h_T}) \geq \gamma^i_{\theta(T)}((a^i, \sigma^i)|_{h_T}), \sigma^{-i}|_{h_T}).$$

$$\big(resp.\ \ \gamma^i_{\theta(T)}(\sigma^i|_{h_T}, \sigma^{-i}|_{h_T}) \geq \gamma^i_{\theta(T)}\big((a^i, \sigma^i)|_{h_T}), \sigma^i|_{h_T}), \sigma^{-i}|_{h_T}) - \varepsilon\big)$$

Remark. The subgame perfect Nash equilibrium for time-inconsistent preferences is different from that in Definition 2.4 for time-inconsistent preferences in two ways. The first is that the evaluation for the subgame at stage T is $\theta(T)$ instead of $\theta[T]$; the second is that we allow for global deviation (coordination at different stages) for time consistency but only one-shot deviation for time inconsistency.

As the preference now involves different evaluations at different stages, a similar but different patience condition is needed.

Definition 6.2 $\{\theta^k(T) : T \geq 1\}_{k\geq 1}$ *is a sequence of **patient** evaluations (time-inconsistent) if:*

$$\forall t, T \in \mathbb{N}^*,\ \lim_{k\to\infty} \frac{\theta^k_t(T)}{\theta^k_{t+1}(T)} = 1, \quad \textit{and the convergence is uniform in both } t \textit{ and } T.$$

Remark As for the special case $\theta(T) = \theta(T')$ for different T, T', the above patience condition reduces to a same form as that in Definition 2.2. Both refer to a single evaluation, but with a different meaning.

Recall that in the proof for the folk theorems with time-consistent preferences, the one-shot deviation principle applied, i.e. checking the optimality immune to global deviation is equivalent to checking that to one-shot deviation. This implies that the proof for folk theorems with time-inconsistent preferences can be carried out in exactly the same line as for time consistency, except for that one has to replace everywhere $\theta[T]$ by $\theta(T)$. Moreover, some of the regularity conditions on $\theta[T]$, such as those proved in Lemma 4.2, are now stated directly by the definition of patience.

Theorem 6.3 ***(Asymptotic Perfect Folk Theorem, time-inconsistent)*** *Let $\{\theta^k(T) : T \geq 1\}_{k\geq 1}$ be a sequence of patient evaluations. Suppose that in a repeated game Γ, int$\mathcal{F}$ is nonempty and fully dimensional, then $E_{\{\theta^k(T)\}} \longrightarrow_{k\to\infty} \mathcal{F}$, where the convergence is with respect to Hausdoff distance.*

The above result can be restated in the following form.

Theorem 6.4 ***(Asymptotic Perfect Folk Theorem restated, time-inconsistent)*** *Suppose that in a repeated game Γ, int$\mathcal{F}$ is nonempty and fully dimensional. Let x be any vector in int$\mathcal{F}$. For any sufficiently small ε, there exists some $\delta_0 \in (0,1)$ such that for all $\{\Lambda(T) : T \geq 1\} = \{(\delta_{T,t})_{t\geq 1} : T \geq 1\}$ with $\delta_{T,t} \in [\delta_0, 1), \forall T, t \geq 1$, then there is some $y \in E_{\{\Lambda(T)\}}$ satisfying $\|x - y\|_\infty \leq \varepsilon$.*

The uniform folk theorems cast in the same way.

Theorem 6.5 ***(Uniform Perfect Folk Theorem, time-inconsistent)*** *The set of perfect uniform equilibrium payoff set coincides with the individual rational set. To be precise: let $\{\theta^k(T) : T \geq 1\}_{k\geq 1}$ be a sequence of patient evaluations and let x be any payoff vector in $\mathcal{F}$. Then for any small $\varepsilon > 0$, there is some behavior strategy $\sigma \in \Sigma$ such that:*

1) $\gamma_{\theta^k(1)}(\sigma) \longrightarrow_{k\to+\infty} y$ *for some* $y \in \mathcal{F}$ *with* $||y-x||_\infty \leq \varepsilon$*;*
2) there exists some $K_0 \in \mathbb{N}$ *satisfying: for all* $k \geq K_0$*,* σ *is an subgame perfect* ε*-Nash equilibrium in the repeated game* $\Gamma_{\{\theta^k(T)\}}$*, i.e.,* $\forall T \geq 1$*,* $\forall h_T \in \mathcal{H}_T$*, and* $\forall i \in I$*,* $\forall a^i \in A^i$*,*

$$\forall i \in I,\ \gamma^i_{\theta^k(T)}(\sigma^i|_{h_T}, \sigma^{-i}|_{h_T}) \geq \gamma^i_{\theta^k(T)}((a^i, \sigma^i)|_{h_T}, \sigma^{-i}|_{h_T}) - \varepsilon.$$

Analogously, we obtain the following equivalent form of the above result.

Theorem 6.6 ***(Uniform Perfect Folk Theorem restated, time-inconsistent)*** *Let* x *be any payoff vector in* $\mathcal{F}$*. Then for any small* $\varepsilon > 0$*, there is some* $\delta_0 \in (0,1)$ *and some behavior strategy* $\sigma \in \Sigma$ *such that: for all* $\{\Lambda(T) : T \geq 1\} = \{(\delta_{T,t})_{t\geq 1} : T \geq 1\}$ *satisfying* $\delta_{T,t} \in [\delta_0, 1), \forall T, t \geq 1$*,*
1) $||\gamma_{\{\Lambda(T)\}}(\sigma) - x||_\infty \leq \varepsilon$*;*
2) σ *is an subgame perfect* ε*-Nash equilibrium in the repeated game* $\Gamma_{\{\Lambda(T)\}}$*, i.e.,* $\forall T \geq 1$*,* $\forall h_T \in \mathcal{H}_T$*, and* $\forall i \in I$*,* $\forall a^i \in A^i$*,*

$$\forall i \in I,\ \gamma^i_{\Lambda(T)}(\sigma^i|_{h_T}, \sigma^{-i}|_{h_T}) \geq \gamma^i_{\Lambda(T)}((a^i, \sigma^i)|_{h_T}, \sigma^{-i}|_{h_T}) - \varepsilon.$$

Finally we point out that the above results in general form imply the following folk theorems for (β, δ)-preference (quasi-hyperbolic discounting), the definition of which is given in Example 2. Let $\theta(\beta, \delta)$ be the associated evaluation with the (β, δ)-preference.

Corollary 6.7 ***(Asymptotic Perfect Folk Theorem with Quasi-hyperbolic Discounting)*** $E_{\theta(\beta,\delta)}$ *converges to* $\mathcal{F}$ *w.r.t. the Hausdoff distance as both* β *and* δ *converge to* 1*.*

Corollary 6.8 ***(Uniform Perfect Folk Theorem with Quasi-hyperbolic Discounting)*** *Let* x *be any payoff vector in* $\mathcal{F}$*. Then for any small* $\varepsilon > 0$*, there is some* $\delta_0 \in (0,1)$ *and some behavior strategy* $\sigma \in \Sigma$ *such that: for any* (β, δ) *satisfying* $\beta, \delta \in [\delta_0, 1)$*,* *1)* $||\gamma_{\theta(\beta,\delta)}(\sigma) - x||_\infty \leq \varepsilon$*;*

2) σ *is an subgame perfect* ε*-Nash equilibrium in all repeated games* $\Gamma_{\theta(\beta,\delta)}$.

7 Appendix

7.1 Proof of Lemma 4.4

Let $x = \sum_{i=1}^{L} \lambda_\ell x_\ell$ be the convex combination for some coefficients (λ_ℓ). W.l.o.g., we may assume each $\lambda_\ell > 0$. Consider $\eta > 0$ with $B_\eta(x) \subseteq co\{x_\ell\}$ and let $\eta' = \eta/5$.

Define a constant $R_{\max}$ as follows. For each ℓ, let $X_\ell = (x_s - x_\ell)_{s \neq \ell}$ be the $I \times (L-1)$ matrix. By assumption $\{x_s - x_\ell : s \neq \ell\}$ are linear independent, so $rank(X_\ell) = L - 1$ thus $(X_\ell X_\ell^T)^{-1}$ exists. We denote $R_\ell = \|X_\ell^T\|_\infty \|(X_\ell X_\ell^T)^{-1}\|_\infty$ and $R_{\max} := \max_\ell\{R_\ell\}$.

Following Lemma 4.2, we take $K(\eta)$ large such that as long as $k \geq K(\eta)$, $\theta_1^k[T] < \frac{1}{L} - 2LR_{\max}\eta'$ for all $T \geq 1$. For fixed $k \geq K(\eta)$ and $\{x_{\ell,m}\} \subseteq B_{\eta'}(x_\ell)$, we construct the sequence $x'_{\varphi(m)}$ as follows.

Decomposition of x into some $x'_{\varphi(1)}$ and some $x_{\phi(1)}$.

Note that $\theta_1^k < \frac{1}{L} - \eta' L R_{\max} < \frac{1}{L} \leq \max\{\lambda_\ell\}$, so there is some $\lambda_{\varphi(1)} \in \{\lambda_1, ..., \lambda_L\}$ with $\lambda_{\varphi(1)} - \theta_1^k > \eta' L R_{\max}$.

Below is a sensitivity result. Set $x'_{\varphi(1)} = x_{\varphi(1),1} \in B_{\eta'}(x_{\varphi(1)})$, then if we replace $x_{\varphi(1)}$ by this slightly perturbed point $x'_{\varphi(1)}$ in the convex combination $x = \lambda_{\varphi(1)} x_{\varphi(1)} + \sum_{\ell \neq \varphi(1)} \lambda_\ell x_\ell$, the coefficient (λ_ℓ)'s perturbation is also small.

Claim 7.1 *Suppose that* x *is expressed by the convex combination* $x = \lambda'_{\varphi(1)} x'_{\varphi(1)} + \sum_{\ell \neq \varphi(1)} \lambda'_\ell x_\ell$ *for some coefficient* (λ'_ℓ)*, then* $|\lambda'_\ell - \lambda_\ell| \leq \eta' L R_{\varphi(1)},\ \forall \ell \in \{1, ..., L\}$.

Proof of Claim 7.1 Since $B_\eta(x) \subseteq co\{x_\ell\}$ and $\eta' < \eta$ the convex combination $x = \lambda'_{\varphi(1)} x'_{\varphi(1)} + \sum_{\ell \neq \varphi(1)} \lambda'_\ell x_\ell$ always exists. Denote $\bar{\epsilon} = x'_{\varphi(1)} - x_{\varphi(1)}$, then $\|\bar{\epsilon}|_\infty \leq \eta'$.

We have

$$x = \lambda_{\varphi(1)} x_{\varphi(1)} + \sum_{\ell \neq \varphi(1)} \lambda_\ell x_\ell = \lambda'_{\varphi(1)} x'_{\varphi(1)} + \sum_{\ell \neq \varphi(1)} \lambda'_\ell x_\ell.$$

Writing $\lambda_{\varphi(1)} = 1 - \sum_{\ell \neq \varphi(1)} \lambda_\ell$ and $\lambda'_{\varphi(1)} = 1 - \sum_{\ell \neq \varphi(1)} \lambda'_\ell$ and arranging terms in the above equation, we obtain

$$\lambda'_{\varphi(1)} \bar{\epsilon} = \lambda'_{\varphi(1)} \big(x'_{\varphi(1)} - x_{\varphi(1)} \big) = \sum_{\ell \neq \varphi(1)} \big(\lambda_\ell - \lambda'_\ell \big) \big(x_\ell - x_{\varphi(1)} \big) \tag{7.1}$$

Recall that $X_{\varphi(1)} = \big(x_\ell - x_{\varphi(1)} \big)_{\ell \neq \varphi(1)}$ is the $I \times (L-1)$ matrix and we denote by $Y_{\varphi(1)} = (\lambda_\ell - \lambda'_\ell)^T_{\ell \neq \varphi(1)}$ the $(L-1) \times 1$ vector, then Equation (7.1) can be written as $X_{\varphi(1)} Y_{\varphi(1)} = \lambda'_{\varphi(1)} \bar{\epsilon}$. Since $L - 1 \leq I$ and $X_{\varphi(1)}$ has rank $L-1$, we deduce that $Y_{\varphi(1)} = X^T_{\varphi(1)} (X_{\varphi(1)} X^T_{\varphi(1)})^{-1} \lambda'_{\varphi(1)} \bar{\epsilon}$ so $\|Y_{\varphi(1)}\|_\infty \leq R_{\varphi(1)} \eta'$, i.e. $|\lambda_\ell - \lambda'_\ell| \leq R_{\varphi(1)} \eta'$ for all $\ell \neq \varphi(1)$. Finally, using the equality $0 = \sum_{\ell=1}^L (\lambda_\ell - \lambda'_\ell) = \lambda_{\varphi(1)} - \lambda'_{\varphi(1)} + \sum_{i \neq \varphi(1)}^L (\lambda_\ell - \lambda'_\ell)$, we obtain that $|\lambda_{\varphi(1)} - \lambda'_{\varphi(1)}| \leq \eta' L R_{\varphi(1)}$. This finishes the proof of the claim. □

Recall that we have chosen $\lambda_{\varphi(1)}$ with $\lambda_{\varphi(1)} - \theta_1^k > \eta' L R_{\max}$. Claim 7.1 applies, giving us $\lambda'_{\varphi(1)} \geq \lambda_{\varphi(1)} - \eta' L R_{\max} > \theta_1^k$. Then it is possible for us to find some $x_{\phi(1)}$ in $co\{x'_{\varphi(1)}, x_\ell | \ell \neq \varphi(1)\}$ such that

$$x = \theta_1^k x'_{\varphi(1)} + (1 - \theta_1^k) x_{\phi(1)}.$$

Indeed, let us just set $x_{\phi(1)} = \frac{1}{1-\theta_1^k} \Big[\big(\lambda'_{\varphi(1)} - \theta_1^k \big) x'_{\varphi(1)} + \sum_{\ell \neq \varphi(1)} \lambda'_\ell x_\ell \Big]$.

Decomposition of $x_{\phi(1)}$ into some $x'_{\varphi(2)}$ and $x_{\phi(2)}$.

Next we focus on $x_{\phi(1)}$ and want to decompose it in the similar way as x. Here the situation is a little bit different because $x_{\phi(1)} \in co\{x'_{\varphi(1)}, x_\ell | \ell \neq \varphi(1)\}$ while $x \in co\{x_\ell\}$, and we need to clarify that this will not make our construction fundamentally different. Especially, the sensitivity result will hold as well and we can allocate the weight $\theta_1^k[2]$ on some $x'_{\varphi(2)}$.

We write $x_{\phi(1)} = \sum_\ell \hat{\lambda}_\ell \hat{x}_\ell$ with the coefficients $\{\hat{\lambda}_\ell\}$ where $\hat{x}_\ell = x_\ell, \forall \ell \neq \varphi(1)$ and $\hat{x}_{\varphi(1)} = x'_{\varphi(1)}$. By dropping unnecessary extreme points, we may always assume that for all $\hat{x}_\ell$, $\{\hat{x}_s - \hat{x}_\ell | s \neq \ell\}$ is linear independent. Define now $\hat{R}_{\max}$ in the same way as $R_{\max}$ but with respect to the points $\{\hat{x}_\ell\}$. Now take k even larger, say $k \geq \max\{\hat{K}(\eta), K(\eta)\}$, where $\hat{K}(\eta)$ is defined in the same way as $K(\eta)$ but with respect to $\hat{R}_{\max}$.

Suppose $\hat{\lambda}_{\varphi(2)} \in \{\hat{\lambda}_\ell\}$ is such that $\hat{\lambda}_{\varphi(2)} - \theta_1^k[2] > 2\eta' L\hat{R}_{\max}$. We fix $x'_{\varphi(2)} := x_{\varphi(2),2}$ that is in $B_{\eta'}(x_{\varphi(2)})$. We would like to write $x_{\phi(1)} = \hat{\lambda}'_{\varphi(2)}x'_{\varphi(2)} + \sum_{\ell\neq\varphi(2)}\hat{\lambda}'_\ell\hat{x}_\ell$ as a perturbation from $x_{\phi(1)} = \hat{\lambda}_{\varphi(2)}\hat{x}_{\varphi(2)} + \sum_{\ell\neq\varphi(2)}\hat{\lambda}_\ell\hat{x}_\ell$. The similar sensitivity result as in Claim 7.1 holds as long as $\|x'_{\varphi(2)} - \hat{x}_{\varphi(2)}\|_\infty$ is small. Indeed, there are two possibilities:

- either $\hat{x}_{\varphi(2)} \neq x'_{\varphi(1)}$ so $\hat{x}_{\varphi(2)} = x_{\varphi(2)}$ and the perturbation $\|x'_{\varphi(2)} - \hat{x}_{\varphi(2)}\|_\infty$ is at most η';

- or $\hat{x}_{\varphi(2)} = x'_{\varphi(1)}$ so $x_{\varphi(2)} = x_{\varphi(1)}$ and the perturbation $\|x'_{\varphi(2)} - \hat{x}_{\varphi(2)}\|_\infty \leq \|x'_{\varphi(2)} - x_{\varphi(1)}\|_\infty + \|x'_{\varphi(1)} - x_{\varphi(1)}\|_\infty$ is at most $2\eta'$ because $x'_{\varphi(2)}$ and $x'_{\varphi(1)}$ are both within η' of $x_{\varphi(1)}$.

Now the same argument as the decomposition of x applies, we can find some $x_{\phi(2)}$ in $co\{\hat{x}_\ell\}$ such that

$$x_{\phi(1)} = \theta_1^k[2]x'_{\varphi(2)} + \big(1 - \theta_1^k[2]\big)x_{\phi(2)}.$$

Then we have

$$x = \theta_1^k x'_{\varphi(1)} + (1-\theta_1^k)\Big\{\theta_1^k[2]x'_{\varphi(2)} + \big(1-\theta_1^k[2]\big)x_{\phi(2)}\Big\} = \theta_1^k x'_{\varphi(1)} + \theta_2^k x'_{\varphi(2)} + \big(1-\theta_1^k-\theta_2^k\big)x_{\phi(2)}.$$

Iteratively, we do the same decompositions for $x_{\phi(2)}, ..., x_{\phi(2)}, ...$ by specifying $x'_{\varphi(3)}, ..., x'_{\varphi(m+1)}, ...$ with

$$x_{\phi(m)} = \theta_1^k[m+1]x'_{\varphi(m+1)} + \big(1 - \theta_1^k[m+1]\big)x_{\phi(m+1)}.$$

In doing this, we obtain $x = \sum_{m\geq1}\theta_m^k x'_{\varphi(m)}$. The last point is to clarify that the parameters $R_{\max}$, $\hat{R}_{\max}, ...,$ thus $K(\eta), \hat{K}(\eta), ...,$ can be taken uniformly bounded. This is true due to the fact that the vectors $\{\hat{x}_\ell\}$ at each step of decomposition are taken within the compact set $\bigcup_{1\leq\ell\leq L} B_{\eta'}(x_\ell)$. This finishes the proof of Lemma 4.4.

7.2 Proof of Proposition 4.3

Fix x in $\mathcal{F}$ and $\eta > 0$. The proof is trivial if $x \in \{g(a) : a \in A\}$. Denote $J := |A|$, and relabel elements in A by $A = \{a(1), ..., a(J)\}$. Set $\eta' = \eta/5$ and let $\mathcal{O}_\eta(x)$ be the open ball of radius η centered at x, relative to $\mathcal{F}$. Take some payoff vectors $x_1, ..., x_L$ in $\mathcal{F}$ such that the following are satisfied:

- x is written as a convex combination of $\{x_\ell : 1 \leq \ell \leq L\}$, that is, $x = \sum_{1\leq \ell \leq L} \alpha_\ell x_\ell$ for the coefficients (α_ℓ) in $[0,1]$ and summing to one;
- $\mathcal{O}_{\eta'}(x) \subseteq co\{x_\ell : 1 \leq \ell \leq L\} \subseteq \mathcal{O}_{2\eta'}(x)$;
- each x_ℓ is a rational convex combination of $\{g(a)\}$, that is, for each ℓ, there are rational coefficients (λ_j^ℓ) in $[0,1]$ summing to one such that $x_\ell = \sum_j \lambda_j^\ell g\big(a(j)\big)$.

According to Carathéodory's theorem, it is possible to have that $L \leq I+1$ and that for each ℓ, the set of vectors $\{x_s - x_\ell : s \neq \ell\}$ is linear independent. Moreover, one can take positive integers (Q_j^ℓ) and Q (which is independent of ℓ) with $\lambda_j^\ell = Q_j^\ell / Q$.

Now we associate with each x_ℓ the cycle of Q action profiles (a *block*): to play $a(1)$ for Q_1^ℓ times, then $a(2)$ for Q_2^ℓ times,..., and finally $a(J)$ for Q_J^ℓ times.

The sequence of pure actions we are going to construct will take the form of one block after another, and each block has length Q and corresponds a different x_ℓ. Let $\mathbf{B}(m) := \{(m-1)Q+1, ..., mQ\}$ be the m-th block. The identification of the sequence $\big(x_{\varphi(m)}\big)$ in $\{x_\ell\}$ will then follows the same procedure as in Lemme 4.4.

For the fixed evaluation $(\theta_t^k)_{t\geq 1}$, we define an associated evaluation $(\beta_m)_{m\geq 1}$ by block: $\beta_m = \sum_{t\in \mathbf{B}(m)} \theta_t^k$ for each $m \geq 1$. k is taken large so $\beta_1[m] \leq \frac{1}{L} - \eta' L R_{\max}$ for all $m \geq 1$.

Define the perturbation vector $x_{\ell,m} \in B_{\eta'}(x_\ell)$ as the θ^k-evaluated average payoff on block $\mathbf{B}(m)$ when x_ℓ is called:

$$x_{\ell,m} = \frac{1}{\beta_m}\left\{ g\big(a(1)\big) \sum_{t\in \mathbf{B}_1^\ell(m)} \theta_t^k + \cdots + g\big(a(J)\big) \sum_{t\in \mathbf{B}_J^\ell(m)} \theta_t^k \right\},$$

where $\mathbf{B}_j^\ell(m)$ is the set of Q_j^ℓ stages playing $a(j)$ within the block $\mathbf{B}(m)$.

We verify that for sufficiently large k, $x_{\ell,m}$ is indeed in $B_{\eta'}(x_\ell)$. According to Lemma 4.2, there is some $K_1(\eta')$ adapted to (Q_j^ℓ) such that: for any $k \geq K_1(\eta')$, $x_\ell^{(m)}$ is η' close to $\frac{Q_1^\ell g\big(a(1)\big)+\cdots+Q_J^\ell g\big(a(J)\big)}{Q}$ (which is actually x_ℓ).

We then use the procedure in Lemma 4.4 to identify the sequence $(x'_{\varphi(m)})$ with each $x'_{\varphi(m)} = x_{\varphi(m),m}$. When $x'_{\varphi(m)}$ is called, the block $\mathbf{B}(m)$ plays the sequence of pure actions corresponding to the vector $x_{\varphi(m),m}$.

Next we show that in doing so, the constructed sequence of pure actions $h_\infty = (a_t)_{t\geq 1}$ satisfy the desired properties.

i). $\gamma_{\theta^k}(h_\infty) = \sum_{t\geq 1}\theta_t^k g\big(a_t\big) = \sum_{m\geq 1}\beta_m x_{\varphi(m),m} = \sum_{m\geq 1}\beta_m x'_{\varphi(m)}$, which is equal to the objective vector x according to Lemma 4.4;

ii). the continuation payoff after any stage T is within $B_\eta(x)$. Indeed, assume that T is within the block $\mathbf{B}(m_0)$ for some m_0. Define $\hat{\mathbf{B}}(m_0) = \mathbf{B}(m_0)\cap\{T+1, ..., \}$. Then by the choice of k being sufficiently large, we have $\xi := \sum_{t\in\hat{\mathbf{B}}(m_0)}\theta_t^k \leq \beta_1[m_0] \leq \frac{\eta'}{M}$. We fix any realization h_∞, and denote by $\mathbf{A}$ the θ^k-evaluation weighted average payoff vector on $\mathbf{B}(m_0)$, then

$$\gamma_{\theta^k}(h_\infty^T) = \eta\mathbf{A} + (1-\eta)\sum_{m\geq m_0+1}\beta_{m-m_0}[m_0]x_{\ell,m},$$

thus

$$\begin{aligned}
\|\gamma_{\theta^k}(h_\infty^T) - x\|_\infty &\leq \xi\|\mathbf{A}-x\|_\infty + \sum_{m\geq m_0+1}\beta_{m-m_0}[m_0]\|x_{\ell,m}-x\|_\infty \\
&\leq 2\eta' + \sum_{m\geq m_0+1}\beta_{m-m_0}[m_0]\Big\{\|x_{\ell,m}-x_\ell\|_\infty + \|x_\ell - x\|_\infty\Big\} \\
&\leq \eta.
\end{aligned}$$

Finally, $K(\eta)$ can be chosen independent of x due to the fact that $\mathcal{F}$ is compact. This finishes the proof of Proposition 4.3.

7.3 Proof of Proposition 5.3

Fix an $\eta > 0$. Let x be a vector in the interior of $\mathcal{F}$. By Carathéodory's theorem, it can be written as a convex combination of at most $(I+1)$ payoff vectors induced by the pure actions profiles in the stage game, i.e. there is some $(\lambda_j)_{1\leq j\leq I+1}$ and $(a^j)_{1\leq j\leq I+1}$ with $x = \sum_j \lambda_j g(a^j)$. We can find (Q_j) some positive integers such that $|Q_j/Q - \lambda_j| \leq \frac{\eta}{2M(I+1)}$ for each $j = 1, ..., I+1$, where $Q = \sum_j Q_j$. Define $y = \sum_j (Q_j/Q) g(a^j)$, then $|y^i - x^i| \leq \eta/2$ for all i.

Now we define cycles of pure actions $(a_t)_{t\geq 1}$: to play a^1 for Q_1 stages,..., a^{I+1} for Q_{I+1} stages, again a^1 for Q_1 stages,... and so on. Let $\mathbf{B}(m) = \{(m-1)Q + 1, ..., mQ\}$ be the m-th bloc of the periodic actions. Let $(\theta^k)_{k\geq 1}$ be a sequence of patient evaluations. According to Lemma 4.1, there is some $K_1(\eta)$ adapted to (Q_j) large such that: for any $k \geq K_1(\eta)$, the weighted average payoff under evaluation θ^k on any block $\mathbf{B}(m)$ is $\eta/2$ close to y. This proves condition $i)$: the total θ-evaluated payoff $z = \sum_{t\geq 1} \theta_t g(a_t)$ is $\eta/2$ close to y, thus η close to x.

Let $K_2(\eta)$ be such that: $\theta_1^k[T] \leq \frac{\eta}{4MQ}$, $\forall T \geq 1, \forall k \geq K_2(\eta)$. We fix any $k \geq K(\eta) := \max\{K_1(\eta), K_2(\eta)\}$, and, to save notation, write θ shortly for θ^k till the end of the proof. Fix any $T \geq 1$ and we suppose that $T \in \mathbf{B}(m)$ for some m. We observe that $\gamma_{\theta[mQ+1]}(h_\infty^{mQ})$ is $\eta/2$ close to y, according to the argument in proving condition $i)$. The payoff $\gamma_{\theta[T+1]}(h_\infty^T)$ is a convex combination of the weighted payoff in stages $\{T+1, ..., mQ\}$ and $\gamma_{\theta[mQ+1]}(h_\infty^{mQ})$, then it is sufficient to bound $\sum_{t=T+1}^{mQ} \theta_{t-T}[T+1]$, the weight of stages $\{T+1, ..., mQ\}$ under $\theta[T+1]$:

$$\left|\gamma_{\theta[T+1]}(h_\infty^T) - \gamma_{\theta[mQ+1]}(h_\infty^{mQ})\right| \leq 2M \sum_{t=T+1}^{mQ} \theta_{t-T}[T+1] \leq 2MQ\theta_1[T+1] \leq \eta/2.$$

This proves condition $ii)$: $|\gamma_{\theta[T+1]}(h_\infty^T) - z| \leq \eta$.

Finally, as in Proposition 4.3, $K(\eta)$ can be chosen independent of x since Δ is compact.

References

[1] I. Arribas and A. Urbano, *Repeated games with probabilistic horizon*, Mathematical Social Sciences, 50: 39–60, 2005.

[2] R. Aumann and S. Shapley, *Long-term competition – A game theoretic analysis*, in Essays on game theory (N. Megiddo, editor), pages 1–15, Springer-Verlag, New-York, 1994.

[3] J.P. Benoit and V. Krishna, *Finitely repeated games*, Econometrica, 53: 905–922, 1985.

[4] J.P. Benoit and V. Krishna, *Nash equilibria of finitely repeated games*, International Journal of Game Theory, 16: 197–204, 1987.

[5] A. Bernergard, *Folk theorem for present-biased players*, SSE/EFI Working Papers Series in Economics and Finance, 2011.

[6] B. D. Bernheim and A. Dasgupta, *Repeated games with asymptotically finite horizons*, Journal of Economic Theory, 67: 129–152, 1995.

[7] H. Chade, P. Prokopovych and L. Smith, *Repeated games with present-biased preference*, Journal of Economic Theory, 139: 157–175, 2008.

[8] S. Frederick, G. Loewenstein and T. O'Donoghue, *Time discounting and time preference: a critical review*, Journal of Economic Literature, 40: 351–401, 2010.

[9] D. Fudenberg and E. Maskin, *The Folk theorem in repeated games with discounting or with incomplete information*, Econometrica, 54: 533–554, 1986.

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

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

[12] O. Gossner *The folk theorem for finitely repeated games with mixed strategies*, International Journal of Game Theory, 24: 95–107, 1995.

[13] A. Kochov, and Y. Song *Repeated games with endogenous discounting*, SSRN Working Paper No.2714337, 2016.

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

[15] D. Laibson, *Golden eggs and hyperbolic discounting*, Quarterly Journal of Economics, 112: 443–77, 1997.

[16] G. Lowenstein and D. Prelec, *Anomalies in intertemporal choice: evidence and an interpretation*, Quarterly Journal of Economics, 107: 573–597, 1992.

[17] G. Mailath and L. Samuelson, *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press, 2006.

[18] J. Renault, *General limit value in dynamic programming*, Journal of Dynamics and Games, 1: 471–484, 2014.

[19] J.F. Mertens, S. Sorin and S. Zamir, *Repeated Games*, Oxford University Press, 2015.

[20] I. Obara and J. Park, *Repeated games with general discounting*, Journal of Economic Theory, 172: 348–375, 2017.

[21] I. Obara and J. Park, *Repeated games with general time preference*, Working paper, 2017.

[22] B. Peleg and M. Yaari, *The existence of a consistent course of action when tastes are changing*, Review of Economic Studies, 40: 391–401, 1973.

[23] A. Rubinstein, *Equilibrium in supergames*, Center for Research in Mathematical Economics and Game Theory, Research Memorandum 25, 1977.

[24] T. Sekiguchi, and K. Wakai, *Repeated games with recursive utility: Cournot duopoly under gain/loss asymmetry*, Kyoto University Discussion Paper No. E-16-006, 2016.

[25] R.H. Strotz, *Myopia and inconsistency in dynamic utility maximization*, Review of Economic Studies, 23: 165–180, 1955-1956.

[26] S. Sorin, *On repeated games with complete information*, Mathematics of Operations Research, 11: 147–160, 1986.