Stochastic Discounting in Repeated Games: Awaiting the Almost Inevitable*

Mehmet Barlo — Sabancı University

Can Ürgün — Sabancı University

January, 2011

Abstract

This paper studies repeated games with pure strategies and stochastic discounting under perfect information. We consider infinite repetitions of any finite normal form game possessing at least one pure Nash action profile. The period interaction realizes a shock in each period, and the cumulative shocks while not affecting period returns, determine the probability of the continuation of the game. We require cumulative shocks to satisfy the following: (1) Markov property; (2) to have a non-negative (across time) covariance matrix; (3) to have bounded increments (across time) and possess a denumerable state space with a rich ergodic subset; (4) there are states of the stochastic process with the resulting stochastic discount factor arbitrarily close to 0, and such states can be reached with positive (yet possibly arbitrarily small) probability in the long run. In our study, a player's discount factor is a mapping from the state space to $(0,1)$ satisfying the martingale property.

In this setting, we, not only establish the (subgame perfect) folk theorem, but also prove the main result of this study: In any equilibrium path, the occurrence of any finite number of consecutive repetitions of the period Nash action profile, must almost surely happen within a finite time window. That is, any equilibrium strategy almost surely contains arbitrary long realizations of consecutive period Nash action profiles.

Journal of Economic Literature Classification Numbers: C72; C73; C79

Keywords: Repeated Games; Stochastic Discounting; Stochastic Games; Folk Theorem; Stopping Time

---

*Any remaining errors are ours.

1 Introduction

The folk theorems of Aumann and Shapley (1994) and Fudenberg and Maskin (1986) establish that payoffs which can be approximated in equilibrium with patient players are equal the set of individually rational ones. The main reason for this observation is players' ability to coordinate their actions using past behavior. In turn, this vast multiplicity of equilibrium payoffs, considerably weakens the predictive power of game theoretic analysis. Moreover, the consideration of limited memory and bounded rationality, lack of perfect observability of the other players' behavior and the past, and uncertainty of future payoffs do not change this conclusion, documented by Kalai and Stanford (1988), Sabourian (1998), Barlo, Carmona, and Sabourian (2009), Barlo, Carmona, and Sabourian (2007); Fudenberg, Levine, and Maskin (1994), Hörner and Olszewski (2006), Mailath and Olszewski (2008); Dutta (1995), Fudenberg and Yamamato (2010), and Hörner, Sugaya, Takahashi, and Vieille (2010). An important aspect of all these findings is the use of constant discounting. The accepted interpretation of the use of discounting in repeated games, offered by Rubinstein (1982) and Osborne and Rubinstein (1994), is that the discount factor determines a player's probability of surviving into the next period. Thus, constant discounting implies that this probability is independent of the history of the game, in particular, invariant.

On the other hand, keeping the same interpretation, but allowing for the discount factor to depend on the history of the game and/or vary across time, is not extensively analyzed in the literature on repeated games. Indeed, to our knowledge, the only relevant work in the study of repeated games is Baye and Jansen (1996) which considers stochastic discounting with period discounting shocks independent from the history of the game. Related work concerning stochastic interest rates can be found in the theory of finance, see Ross (1976), Harrison and Kreps (1979), and Hansen and Richard (1987).

This paper studies repeated games with pure strategies and common stochastic discounting under perfect information. We consider infinite repetitions of any finite normal form game possessing at least one pure Nash action profile. The period interaction realizes a shock in each period, and the cumulative shocks while not affecting period returns, determine the probability of the continuation of the game. We require cumulative shocks to satisfy the following: (1) Markov property; (2) to have a non-negative (across time) covariance matrix; (3) to have bounded increments (across time) and possess a denumerable state space with a rich ergodic subset; (4) there are states of the stochastic process with the resulting stochastic discount factor arbitrarily close to 0, and such states can be reached with positive (yet possibly arbitrarily small) probability in the long run. In our study, a player's discount factor is a monotone mapping from the state space to $(0, 1)$ satisfying the martingale property.

In this setting, we, not only establish the (subgame perfect) folk theorem, but also prove the main result of this study: Under any subgame perfect equilibrium strategy, the occurrence of any finite number of consecutive repetitions of the period Nash action profile, must almost surely happen within a finite time window. That is, any equilibrium strategy almost surely contains arbitrary long realizations of consecutive period Nash action profiles. In other words, every equilibrium outcome path almost surely involves a stage, i.e. the stochastic process governing the one–shot discount factor possesses a *stopping time*, after which long consecutive repetitions of the period Nash action profile must be observed.

Considering the repeated prisoners' dilemma with pure strategies and stochastic discounting, our results display that: (1) the subgame perfect folk theorem holds; and, (2) in any subgame perfect equilibrium strategy for any natural number K, the occurrence of K consecutive defection action profiles must happen almost surely within a finite time period.

An important implication of our main result concerns social contracts and institutions. It is well known that in the study of repeated games with patient players social contracts and institutions, in general, do not create additional equilibrium opportunities to players even under considerations of limited memory and bounded rationality, lack of perfect observability of the other players' behavior and the past, and uncertainty of future payoffs. Our main result, on the other hand, implies that no matter what the initial equilibrium arrangement is, almost surely there will be a period after which the contributions of social contracts and institutions may be positive.

The fundamental reason of our main result is captured by a significant phrase to be found on page 101 of Willams (1991): "Whatever always stands a reasonable chance of happening, will almost surely happen – sooner rather than later." Indeed, due to the restrictions on the stochastic processes we prove that for any $\varepsilon > 0$, the one–shot discount factor must almost surely fall below ε in a finite time period. This result still holds even when the additional requirement of mean-reversion (in the limit) of the one–shot discount rates is employed. Then, given any natural number K, the restriction of bounded increments enable us to identify the level of ε (via the use of K) so that: In any equilibrium path, the one–shot discount factors cannot exceed a certain threshold even when $K + 1$ consecutive "good" shocks are realized. Hence, the occurrence of K consecutive repetitions of the period Nash action profile, must almost surely happen within a finite time window under any subgame perfect strategy.

In order to see why the subgame perfect folk theorem holds, first, notice that due to restricting attention to perfect information and stochastic processes with the Markov property, given any history of shocks, players evaluate future payoffs with their expected discount factors and the conclusions of Abreu (1988) and Sugaya (2010) apply. [1] Moreover, we show that the following observation holds regarding players' expectations for future discount factors: In any period t with any given history of shocks up to that period, each player evaluates future returns at least as much as a player using a constant discount factor obtained from the same shocks. That is, each player's expectation of the discount factor from period t into

[1] Given any t and any history of shocks up to period t, the set of subgame perfect continuation payoffs is compact with players' expected discount factors obtained from these shocks.

period τ, $\tau > t$, is not less than the discount factor from t into $t+1$ raised to the power of $\tau - (t+1)$. Hence, one may approximate a given strictly individually rational payoff vector by constructing a simple strategy profile (supporting that payoff vector via period–0 expectations) involving optimal penal codes to be identified for any given history of shocks.

The literature on stochastic discounting in repeated games is surprisingly not very rich. A significant contribution in that field is Baye and Jansen (1996). This study considers a form of stochastic discounting with no stringent restrictions on the values that one–shot discount factor can take, and, the distribution of one–shot discount factors may depend on the time index. However, such a distribution in a particular period is independent from the past distributions. Moreover, they identify two significant cases: The first, when the one–shot discount factor is realized before the actions in the stage game are undertaken; the second, when the actions need to be chosen before the one–shot discount factor is realized. They prove that the folk theorem holds with in the latter case. They also establish that in the first case, "a full folk theorem is unobtainable... since average payoffs on the efficiency frontier are unobtainable as Nash equilibrium super–game payoffs".

Our paper deals with history dependent stochastic discounting, and the situation when the actions need to be chosen before the one–shot discount factor is realized. That is why, the results regarding the folk theorem presented in Baye and Jansen (1996), is not in contrast with ours. It also needs to be emphasized that, while their restrictions on the one–shot stochastic discount rate are not as strong as ours comparing these distributions across time periods, ours involve the critical aspect of history dependence. Another important point we wish to remind the reader of, is that the main message of this study is rather not the folk theorem.

There is a number of notable contributions in the context of stochastic games. Indeed, recent studies by Fudenberg and Yamamato (2010) and Hörner, Sugaya, Takahashi, and Vieille (2010) generalize the folk theorem of Dutta (1995) for irreducible stochastic games with the requirement of a finite state space. Our setup can be expressed as an irreducible stochastic game where the each players' discounting is constant, yet their payoffs are all obtained from a (stochastic) scalar, and the actions chosen have no bearing on the future payoffs. Indeed, ours is a particular irreducible stochastic game with a denumerable state spaces, hence these folk theorems do not apply.

The organization of the paper is as follows: The next section will present the basic model, notation and definitions, some preliminary yet important results. In section 3 we present the main Theorem of this paper, and section 4 contains the folk theorem. Finally, in section 5, we will present our work solving the repeated prisoners' dilemma with stochastic discounting.

2 Notations and Definitions

Let $G = (N, (A_i, u_i)_{i \in N})$ be a normal form game with $|N|, |A_i| \in \mathbb{N}$ for all $i \in N$; and i's payoff function denoted by $u_i : A \to \mathbb{R}$ where $A = \prod_{i \in N} A_i$ and $A_{-i} = \prod_{j \neq i} A_i$. We let $\Sigma_i = \Delta(A_i)$ where Δ is the simplex and $\Sigma = \prod_{i \in N} \Sigma_i$ and $\Sigma_{-i} = \prod_{j \neq i} \Sigma_i$. $u_i : \Sigma \to \mathbb{R}$ denotes the usual mixed strategy extension of u_i. For any $i \in N$ denote respectively the *minmax payoff* and a *minmax profile* for player i by $v_i = \min_{\sigma_{-i} \in \Sigma_{-i}} \max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \sigma_{-i})$ and $m^i \in \arg\min_{\sigma_{-i} \in \Sigma_{-i}} \max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \sigma_{-i})$.

Let $\mathcal{U} = \{u \in \mathrm{co}\,(u(A)) : u_i \geq v_i \text{ for all } i \in N\}$ denote the set of *individually rational payoffs* and $\mathcal{U}^0 = \{u \in \mathrm{co}(u(A) : u_i > v_i \text{ for all } i \in N\}$ denote the set of *strictly individually rational payoffs*. The game G is *full-dimensional* if the interior of $\mathcal{U}$ in $\mathbb{R}^n$ is nonempty.

For what follows, we assume that G has a pure strategy Nash equilibrium:

Assumption 1 *There exists $a^* \in A$ with $u_i(a^*) \geq u_i(a_i, a^*_{-i})$ for all $i \in N$.*

The *supergame of* G consists of an infinite sequence of repetitions of G taking place in periods $t = 0, 1, 2, 3, \ldots$. Let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

In every period $t \in \mathbb{N}_0$, the probability of the game continuing into the next period is determined by a random variable $\mathbf{X}_t$. The following Assumption needed, allows for a wide class of random variables:

Assumption 2 *For all $t \in \mathbb{N}_0$ random variable $\mathbf{X}_t$ has a countable non-empty and non-trivial support in $\mathbb{R}$ which is bounded from at least one side, and its expectation exists. Moreover, $\mathbf{X}_0 = d \in \mathbb{R}$, and $\mathbf{X}_t \stackrel{d}{=} \mathbf{X}_\tau$ for all $t, \tau \in \mathbb{N}$.*

Let $\mathbf{X} \stackrel{d}{=} \mathbf{X}_t$ for all $t \in \mathbb{N}$, and $X_t \in \mathbb{R}$ denote a particular realization of $\mathbf{X}$ at time t. The aggregation of previous shocks will be done as follows:

Assumption 3 *$\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$ is a stochastic process satisfying the following:*

1. *the Markov property,*
2. *its transition is governed monotonically by $\mathbf{X}$ in a bounded fashion as follows: This transition is strictly increasing (strictly decreasing) whenever the support of $\mathbf{X}$ is bounded from below (above); and is either strictly increasing or strictly decreasing if the support of $\mathbf{X}$ is bounded.*
3. *it has a non-negative covariance matrix across time, whose entries strictly decrease as the relevant time differences strictly increase.*
4. *given the state space, S, of $\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$, the set of ergodic states, denoted by E, is dense in S.*

Without loss of generality assume that the support of $\mathbf{X}$ is bounded above. Then, Assumptions 2 and 3 imply that (1) $\mathbf{Y}_0$ is deterministic, and in particular $\mathbf{Y}_0 = \bar{\nu}(d)$, $\bar{\nu} : \mathbb{R} \to \mathbb{R}$ strictly decreasing; and due to the Markov property (2) $\mathbf{Y}_{t+1} = \nu\left(\mathbf{Y}_t, \mathbf{X}\right)$ where $\nu : \mathbb{R}^2 \to \mathbb{R}$ is strictly decreasing and measurable in the σ-algebra where $\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$ is defined, and bounded from above; and (3) covariance of $\mathbf{Y}_t$ and $\mathbf{Y}_\tau$, $\tau > t$, is greater or equal to 0, and is strictly decreasing as $\tau - t$ strictly increases; and (4) the set of aperiodic and non-transient states of $\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$ must be dense in the state space. Thus, by restricting attention to stochastic processes with the Markov property, this assumption rules out some interesting situations such as the following: Given t, let $\mathbf{Y}_t = \sum_{k=0}^{t} \alpha^{t-k}\mathbf{X}$, $\alpha \in (0,1)$, a formulation in which the more recent shocks are more important then those that occurred earlier.

Given a stochastic process $\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$, let $\{\mathcal{F}_t\}_{t\in\mathbb{N}_0}$ denote its natural filtration (i.e. sequence of growing σ-algebras); and for any given $t \in \mathbb{N}_0$, $\mathcal{F}_t$ is commonly interpreted as the information in period t.

Given t, we let a particular realization of the stochastic process $\{\mathbf{Y}_t\}_{t\in\mathbb{N}}$ be denoted by $Y_t \in \mathbb{R}$.

The supergame is defined for a given $\mathbf{X}_0 = d \in \mathbb{R}$, and is denoted by $G(d)$. [2] For $k \geq 1$, a $k-$*stage history* is a k-length sequence $h_k = ((a_0, X_1), \ldots, (a_{k-1}, X_k))$, where, for all $0 \leq t \leq k-1$, $a_t \in A$; and for all $1 \leq t \leq k$, X_k is realization of $\mathbf{X}$ at time k; the space of all k-length histories is H_k, i.e., $H_k = (A \times \mathbb{R})^k$. We use e for the unique 0–stage history — it is a 0–length history that represents the beginning of the supergame. The set of all histories is defined by $H = \bigcup_{n=0}^{\infty} H_n$. For every $h \in H$ we let $\ell(h)$ denote the *length of* h. For $t \geq 2$, we let $X^t = (X_1, \ldots, X_t)$ denote the history of shocks up to and including period t.

We assume that players have *complete information*. That is, in period $t > 0$, knowing the history up to period t, given by h_t, the players make simultaneous moves denoted by $a_{t,i} \in A_i$. The players' choices in the unique 0–length history e are in A as well. Notice that in our setting, given t, a player not only observes all the previous action profiles, but also all the shocks including the ones realized in period t. In other words, the period–t shocks are commonly observed before making a choice in period t.

[2] Notice that our formulation is similar to the convention of defining repeated games for a constant and common discount factor, often denoted by $G(\delta)$, where $\delta \in (0,1)$ stands for the discount factor.

For all $i \in N$, a *strategy* for player i is a function $f_i : H \rightarrow A_i$ mapping histories into actions. The set of player i's strategies is denoted by F_i, and $F = \prod_{i \in N} F_i$ is the joint strategy space. Finally, a strategy vector is $f = (f_1, \ldots, f_n)$. Given an individual strategy $f_i \in F_i$ and a history $h \in H$ we denote the *individual strategy induced at* h by $f_i|h$. This strategy is defined point-wise on H: $(f_i|h)(\bar{h}) = f_i(h \cdot \bar{h})$, for every $\bar{h} \in H$. We will use $(f|h)$ to denote $(f_1|h, \ldots, f_n|h)$ for every $f \in F$ and $h \in H$. We let $F_i(f_i) = \{f_i|h : h \in H\}$ and $F(f) = \{f|h : h \in H\}$.

A strategy $f \in F$ induces an outcome $\pi(f)$ as follows: $\pi^0(f) = f(e) \in A$; and for $X_1 \in \mathbb{R}$ we have $\pi^1(f)(X^1) = f(f(e), X_1) \in A$; and, $\pi^2(f)(X^2) = f(f(e), f(f(e), X_1), X_2) \in A$, $X_1, X_2 \in \mathbb{R}$; and continuing in this fashion we obtain

$$\pi^k(f)(X^k) = f\left(\pi^0(f), \pi^1(f)(X^1), \ldots, \pi^{k-1}(f)(X^{k-1}), X_k\right) \in A, k > 1 \text{ and } X_1, \ldots, X_k \in \mathbb{R}.$$

2.1 Players' Payoffs

In this subsection, we wish to present the construction of expected payoffs. Due to that regard, first we will present our stochastic discounting construction, and second formulate the resulting expected utilities.

Player i's discounting is stochastic: The stochastic discount factor of player i is a random variable, denoted by $\left\{\mathbf{d}_t^{t+1}\right\}_{t \in \mathbb{N}_0}$, where for any given $t \in \mathbb{N}_0$, $\mathbf{d}_t^{t+1}$ identifies the common probability of the game continuing from period t to period $t+1$.

The particular fashion in which we relate shocks to stochastic discounting is captured in the following assumption:

Assumption 4 *For any $i \in N$, the stochastic discounting process of player i, $\left\{\mathbf{d}_t^{t+1}\right\}_{t \in \mathbb{N}_0}$, satisfies the following:*

1. *Given any $t \in \mathbb{N}_0$, $\mathbf{d}_t^{t+1} = \psi \circ \mathbf{Y}_t$ where ψ is a strictly increasing (strictly decreasing) and continuous function with range given by $(0, 1)$ whenever the support of $\mathbf{X}$ is bounded*

above (below). ψ is either strictly increasing or strictly decreasing whenever the support of $\mathbf{X}$ *is bounded.*

2. $\{\mathbf{Y}_t\}_{t\in\mathbb{N}_0}$ *and* ψ *are such that* $\{\psi \circ \mathbf{Y}_t\}_{t\in\mathbb{N}_0}$ *is a martingale.*

3. $\{\mathbf{Y}_t\}_{t\in\mathbb{N}_0}$ *and* ψ *are such that for any* $t \in \mathbb{N}_0$ *and for any* $\varepsilon > 0$ *there exists* $\tau \geq t$ *with* $\Pr(\psi \circ \mathbf{Y}_\tau < \varepsilon \mid \mathcal{F}_t) > 0$.

For what follows assume without loss of generality that the support of $\mathbf{X}$ is bounded above. Indeed, this situation corresponds to higher shocks to be interpreted as "positive" with respect to the probability of the continuation of the game, because such shocks will result in an increase in the current discount factor.

Then, the above assumption involves the following: First, the discount rate is obtained from the underlying stochastic process determining the cumulative shocks, namely $\{\mathbf{Y}_t\}_{t\in\mathbb{N}_0}$, by employing a continuous and strictly increasing transformation ψ with range $(0,1)$. In fact, this implies that the stochastic process determining the discount rates is bounded.

Moreover, this part of the Assumption also binds the particular *timing* and *information setting* that we employ: Given $\mathbf{X}_0 = d$ (implying $\mathbf{Y}_0 = \bar{\nu}(d)$, the stochastic discount factor determining the probability that the game continues into the next period, $\mathbf{d}_0^1$, is pinned down to a constant $\psi(\bar{\nu}(d))$. In the next period, $t = 1$, $\mathbf{X}_1$ is realized before players decide on $a_1 \in A$. So, $\mathbf{Y}_0 = \bar{\nu}(d)$ and the realization of $\mathbf{X}_1$ together determine (via $\bar{\nu}$) a realization of Y_1, hence, $\delta_1^2 = \psi(\nu(\bar{\nu}(d), X_1))$. Thus, following an inductive argument in any period $t > 1$, the given X^t pins down Y_t which in turn determines the particular level of δ_t^{t+1}, i.e. the probability that the game continues from period t into period $t+1$.

Therefore, our formulation involves δ_t^{t+1} being common knowledge among the players in period t before a_t is chosen. Thus, apart from the beginning of the game our formulation corresponds to the case in Baye and Jansen (1996) where "... players choose actions in each period after having observed the current discount factor". They show that then the Folk

Theorem "...breaks down; payoffs on the boundary of the set of individually rational payoffs are unobtainable as Nash equilibrium average payoffs to the supergame." However, it is important to emphasize the following: (1) Their stochastic discounting formulation involves a common discount factor determined by a random variable distributed independently from the history of the game; and (2) While our formulation necessitates (due to the use of stochastic processes) the period 0 discount factor to be deterministic, the failure of the Folk Theorem shown in the setting of Baye and Jansen (1996) is primarily due to the action profile chosen in period 0 being a function of the random period 0 discount factor (drawn before the period 0 action is chosen).

The second part of Assumption 4 requires the stochastic discounting process to be a martingale. In fact, due to $\{\mathbf{Y}_t\}$ satisfying the Markov property (by condition 1 of Assumption 3), now $\{\mathbf{d}_t\}$ is a martingale with the Markov property. Moreover, because the unconditional expectations of $\psi \circ \mathbf{Y}_t$, $t \in \mathbb{N}_0$, are equal to the conditional expectations with respect to the trivial σ-algebra, $\mathcal{F}_0$, the unconditional expectations do not depend on the time index. This, in turn, implies that $\mathrm{E}(\psi \circ \mathbf{Y}_t) = \mathrm{E}(\psi \circ \mathbf{Y}_\tau) = \mathrm{E}\left(\psi \circ \mathbf{Y}_t | \mathcal{F}_0\right) = \mathrm{E}\left(\psi \circ \mathbf{Y}_\tau | \mathcal{F}_0\right)$ for all $t, \tau \in \mathbb{N}_0$. Thus, the expected value of the discount factor formed at the beginning of the game, is required to be constant.

Finally, in the third part of Assumption 4 we require that there are states of the stochastic process with the resulting stochastic discount factor arbitrarily close to 0, and such states can be reached with positive (yet possibly arbitrarily small) probability in the long run.

An important implication of this Assumption concerns the situation where the common stochastic discount factor is determined by a stochastic process with a finite state space in $(0, 1)$. Then, the third part of Assumption 4 cannot hold. [3] Thus, our analysis excludes

[3]On the other hand, if given a finite state space $S \subset [0, 1)$ with $Y(s) = 0$ for some s and $S = (s_1, \ldots, s_K)$, $K \in \mathbb{N}$, without loss of generality we may assume that $Y(s_K) \geq Y(s)$ for all $s \in S$; thence, the monotonicity requirement in the second part of Assumption 3 cannot hold. Because, due to the (strict) monotonicity requirement, there has to be a $s' \in S$ with $Y(s') > Y(s_K)$. Indeed, this implies that S has to be denumerable.

stochastic processes with finite state spaces.

The stochastic discount factor from period t to period τ, with $\tau \geq t$, given $\mathcal{F}_s$, $s \leq t$, is defined by $\mathbf{d}_t^\tau = \prod_{s=t}^{\tau-1} \mathbf{d}_s^{s+1}$, with the convention that $\mathbf{d}_t^t = 1$. We denote $\mathrm{E}(\mathbf{d}_t^{t+1} \mid \mathcal{F}_s)$ for $s \leq t-1$, by $\mathrm{E}_s\left(\mathbf{d}_t^{t+1}\right)$, which indeed is the projection of $\mathbf{d}_t^{t+1}$ on $\mathcal{F}_s$. For any $t \in \mathbb{N}_0$, we let a realization of $\mathbf{d}_t^{t+1}$ be denoted by δ_t^{t+1}, which stands for the realized probability that the game continues from period t to period $t+1$.

The following Lemma display that the stochastic discounting process constructed in this study involves weaker discounting than the one associated with constant discounting:

Lemma 1 *Suppose that Assumptions 2, 3, and 4 are satisfied. Then*

1. *every possible realization of* $\mathbf{d}_t^\tau$ *is in* $(0,1)$ *for every* $\tau, t \in \mathbb{N}_0$ *with* $\tau > t$, *and the set of states of* $\{\mathbf{d}_t^{t+1}\}_{t \in \mathbb{N}_0}$ *is denumerable,*
2. $\mathrm{E}\left(\mathbf{d}_t^{t+1} | \mathcal{F}_0\right) = \delta_{(0)}$ *for some* $\delta_{(0)} \in (0,1)$ *and for all* $t \in \mathbb{N}_0$,
3. *for every given* $\hat{\delta} \in (0,1)$, *there exists* $d \in \mathbb{R}$ *such that* $\delta_{(0)} = \hat{\delta}$,
4. *for every* $\tau, t, s \in \mathbb{N}_0$ *with* $\tau > t \geq s$, *given* $\mathbf{d}_t^{t+1} = \delta_t^{t+1}$

$$\mathrm{E}\left(\mathbf{d}_\tau^{\tau+1} | \mathcal{F}_t\right) = \delta_t^{t+1}, \tag{1}$$

and

$$\mathrm{E}\left(\mathbf{d}_{t+1}^{\tau} | \mathcal{F}_s\right) \geq \left(\mathrm{E}\left(\mathbf{d}_{t+1}^{t+2} | \mathcal{F}_s\right)\right)^{\tau-(t+1)}. \tag{2}$$

The implications of this Lemma are essential for the proof, the interpretation and the evaluation of our results:

The first one displays that the stochastic process specified has a denumerable (countable but not finite) state space and results in a well-defined construction for stochastic discounting. This is because, $\mathbf{X}$ has a countable support and for every $\tau, t \in \mathbb{N}_0$ with $\tau > t$, $\mathbf{d}_t^\tau = \prod_{s=t}^{\tau-1} \mathbf{d}_s^{s+1}$ and every possible realization of $\mathbf{d}_s^{s+1}$ is in $(0,1)$ for every $s \in \mathbb{N}_0$.

The second shows that date zero expectations of future one–period discount factors are constant with respect to the time index. And the third, displays that d can be chosen so that any given constant discount factor can be precisely obtained. The reason is: Both $\mathbf{Y}_0$ and $\mathbf{X}_0$ are deterministic and recall that $\mathbf{X}_0 = d$ and $\mathbf{Y}_0 = \bar{\nu}(d)$. So, one may select d appropriately (which in turn pins down ψ by Assumption 4) such that $\delta_{(0)} = \psi((\bar{\nu}(d)) = \hat{\delta}$. Therefore, we conclude with respect to date zero expectations the repeated game at hand can be associated with one having a constant and common discount factor. Thus, our repeated game with stochastic discounting can be interpreted as a perturbation of a "standard" repeated game under perfect information with common and constant discount factor (given by $\hat{\delta}$).

Finally, the fourth implication is twofold: Given any history of shocks up to time period t, the first is that the expected level of future one–period discount factors are equal to the current one. The second shows that every player values future returns more than a player using a constant discount factor obtained from the same shocks. That is, a player discounts a return in period τ, $\tau > t$, with $\mathrm{E}_t\left(\mathbf{d}_t^\tau\right)$ which is greater or equal to $\left(\mathrm{E}_t\left(\mathbf{d}_t^{t+1}\right)\right)^{\tau-t}$. (Notice that given X^t, $\mathrm{E}_t\left(\mathbf{d}_t^\tau\right) = \delta_t^{t+1}\mathrm{E}_t\left(\mathbf{d}_{t+1}^\tau\right)$, because $\mathbf{d}_t^{t+1} = \delta_t^{t+1}$ is realized.) In particular, this implies d can be chosen so that

$$\mathrm{E}_0\left(\mathbf{d}_t^\tau\right) \geq \left(\mathrm{E}_0\left(\mathbf{d}_t^{t+1}\right)\right)^{\tau-t} = \left(\delta_{(0)}\right)^{\tau-t} = \hat{\delta}^{\tau-t},$$

and when $\tau = t + 1$ then this inequality holds with an equality. Hence, these properties establish that with a date 0 point of view our stochastic discounting construction involve weaker discounting than that associated with a constant and common discount factor $\hat{\delta}$.

The following Remark summarizes these observations:

Remark 1 *Suppose that Assumptions 2, 3, and 4 hold. Then, given any repeated game under perfect information and a common and constant discount factor* $\hat{\delta} \in (0, 1)$*, there exists a repeated game under perfect information and stochastic discounting specified as above which* *satisfies the following properties: (1) The date zero expectations of the one–shot discount factors are all equal to $\hat{\delta}$; and (2) in date 0 players employ weaker discounting than that associated with a constant and common discount factor $\hat{\delta}$; and (3) expected level of future one–period discount factors are equal to the current one.*

Proof of Lemma 1. 1, 2 and 3 of the Lemma are straightforward as discussed above. The first part of the fourth result is, in fact, the martingale identity. For the second part, notice that due to part 1 of the the fourth result, we can write the conditional expectation $\mathrm{E}\left(\mathbf{d}_{t+1}^{\tau}|\mathcal{F}_t\right)$ as

$$\mathrm{E}\left(\prod_{s=t+1}^{\tau}\mathbf{d}_s^{s+1}|\mathcal{F}_t\right)=\prod_{s=t+1}^{\tau}\mathrm{E}\left(\mathbf{d}_s^{s+1}|\mathcal{F}_t\right)+\Sigma_t^{\tau}=\left(\mathrm{E}\left(\mathbf{d}_{t+1}^{t+2}|\mathcal{F}_t\right)\right)^{\tau-(t+1)}+\Sigma_t^{\tau}.$$

The proof concludes because by Assumptions 3 and 4 Σ_t^{τ} is non-negative. ■

The next Assumption is about players' knowledge of the past when taking expectations:

Assumption 5 *In every period $t \in \mathbb{N}_0$, each player uses the most up to date information, i.e. $\mathcal{F}_t$.*

Given a strategy profile f, because that each period's supremum return is bounded for every player, the payoff of player $i \in N$ in the supergame $G(d)$ of G is, where $\psi(\bar{\nu}(d)) = \hat{\delta} \in (0,1)$:

$$\begin{aligned}U_i(f) \;=\;& (1-\hat{\delta})u_1\left(\pi^0(f)\right)\\ &+(1-\hat{\delta})\mathrm{E}\left(\delta_0^1 u_1\left(\pi^1(f)(X^1)\right)|\mathcal{F}_0\right)\\ &+(1-\hat{\delta})\mathrm{E}\left(\mathrm{E}\left(\delta_0^2 u_1\left(\pi^2(f)(X^2)\right)|\mathcal{F}_1\right)|\mathcal{F}_0\right)\\ &+(1-\hat{\delta})\mathrm{E}\left(\mathrm{E}\left(\mathrm{E}\left(\delta_0^3 u_2\left(\pi^3(f)(X^3)\right)\right)|\mathcal{F}_2\right)|\mathcal{F}_1\right)|\mathcal{F}_0\right)+\ldots.\end{aligned}$$

Because that $\{\mathcal{F}_s\}_{s=0,1,,2,\ldots}$ is the natural filtration, the above term reduces to

$$\begin{aligned} U_i(f) &= (1-\hat{\delta})u_1\left(\pi^0(f)\right) \\ &+(1-\hat{\delta})\mathrm{E}\left(\delta_0^1 u_1\left(\pi^1(f)(X^1)\right)|\mathcal{F}_0\right) \\ &+(1-\hat{\delta})\mathrm{E}\left(\delta_0^2 u_1\left(\pi^2(f)(X^2)\right)|\mathcal{F}_0\right) \\ &+(1-\hat{\delta})\mathrm{E}\left(\delta_0^3 u_1\left(\pi^3(f)(X^3)\right)|\mathcal{F}_0\right) + \ldots, \end{aligned}$$

i.e.

$$U_i(f) = (1-\hat{\delta})\sum_{k=0}^{\infty}\mathrm{E}\left(\delta_0^k u_1\left(\pi^k(f)(X^k)\right)|\mathcal{F}_0\right), \tag{3}$$

where $\pi^0(f)(X^0) = \pi(f(e))$, and recall that $\mathrm{E}\left(\delta_t^t|\mathcal{F}_s\right) = 1$ for all $s \leq t$. Following a similar method, we can also define the continuation utility of player i as follows: Given $t \in \mathbb{N}$ and $X^t \in \mathbb{R}^t$

$$V_i^{t,X^t}(f) = (1-\hat{\delta})\sum_{k=t}^{\infty}\mathrm{E}\left(\delta_t^k u_i\left(\pi^k(f)(X^k)\right)|\mathcal{F}_t\right), \tag{4}$$

and for $\tau > t$

$$V_i^{\tau,X^t}(f) = (1-\hat{\delta})\sum_{k=\tau}^{\infty}\mathrm{E}\left(\delta_\tau^k u_i\left(\pi^k(f)(X^k)\right)|\mathcal{F}_t\right). \tag{5}$$

We use the convention that $V_i^{0,X^0}(f) = U_i(f)$.

2.2 Subgame Perfect Equilibria

A strategy vector $f \in F$ is a *Nash equilibrium* of the supergame of G if for all $i \in N$, $U_i(f) \geq U_i(\hat{f}_i, f_{-i})$ for all $\hat{f}_i \in F_i$. A strategy vector $f \in F$ is a *subgame perfect equilibrium* of the supergame $G(d)$ if every $\bar{f} \in F(f)$ is a Nash equilibrium.

For any $\hat{\delta} \in (0,1)$, let $\mathcal{V}(\hat{\delta})$ be the set of subgame perfect equilibrium payoffs. Moreover, $\mathcal{V}(X^t)$ are the subgame perfect equilibrium continuation payoffs (in period t terms), when X^t is realized. In fact, abusing notation we let $\mathcal{V}(X^t) = \mathcal{V}(\delta_t^{t+1})$. Below we will show that

$\mathcal{V}(\delta) \subseteq \mathcal{U}$ is compact for all $\delta \in (0,1)$, hence obtain the following characterization analogous to Abreu (1988): A strategy f is subgame perfect if and only if for all $i \in N$ and for all $t \in \mathbb{N}_0$ and for all $X^t \in \mathbb{R}^t$, we have

$$V_i^{t,X^t}(f) \geq (1-\hat{\delta}) \max_{a_i \in A_i} u_i(a_i, \pi_{-i}^t(f)(X^t)) + \delta_t^{t+1} \mathrm{E}\left(v_i(X^{t+1})|\mathcal{F}_t\right), \tag{6}$$

where $\delta_t^{t+1} = \mathbf{d}_t^{t+1}$ (i.e. given X^t, the realization of $\mathbf{d}_t^{t+1}$ is equal to δ_t^{t+1}), and for every $i \in N$

$$v_i(X^{t+1}) = \min\left\{u_i : u_i \in \mathcal{V}(\delta_{t+1}^{t+2})\right\}. \tag{7}$$

Before the justification of these, we wish to describe the resulting construction briefly. Notice that, when player i decides whether or not to follow the equilibrium behavior in period t given the history of shocks X^t, it must be that: The player i's expected continuation payoff associated with the equilibrium behavior must be as high as player i deviating singly and optimally today, and being punished tomorrow. An important issue to notice is that, tomorrow players will know X_{t+1} before deciding on their actions. Thus, players will be punishing player i, the deviator, with the most severe and credible punishment with the information they have in period $t+1$. Thus, the punishment payoff to player i with the information that players have in period $t+1$, i.e. X^{t+1}, is $v_i(X^{t+1})$. Player i forecasts these in period t, and hence, forms an expectation regarding his punishment payoff (starting from period $t+1$ onwards) with the information that he has in period t, namely X^t.

In order to show that for every t and X^t $\mathcal{V}(\delta_t^{t+1})$ is compact, we will be employing the construction of Abreu, Pearce, and Stachetti (1990), and it is important to point out that their assumptions, 1-5 are all satisfied in our framework: A2, A3, and A4 are trivially satisfied as the period payoffs are deterministic, and we also impose A1 and A5.

Following their construction, given X^t for any $W \subset \mathbb{R}^N$ and the resulting level of $\delta_t^{t+1} \in (0,1)$, let $g(\delta_t^{t+1})$ be the expected discounted, continuation utility level (not including today's payoff levels and using the normalization via $\hat{\delta} \in (0,1)$), for an arbitrary strategy profile. Furthermore, for that given level of δ_t^{t+1}, consider the pair, $(g(\delta_t^{t+1}), a)$ and define $\mathrm{E}(g(\delta_t^{t+1}), a) = \delta_t^{t+1}\left((1-\hat{\delta})u(a) + g(\delta_{t+1}^{t+2})\right)$. A pair $(g(\delta_t^{t+1}), a)$ is called *admissible with respect to* W if $\mathrm{E}_i(g(\delta_t^{t+1}), a) \geq \mathrm{E}_i\left(g(\delta_t^{t+1}), (\gamma_i, a_{-i})\right)$ for all $\gamma_i \in A_i$ and for all $i \in N$. Moreover, for each set W, define $B^{X^t}(W) = \{\mathrm{E}(g(\delta_t^{t+1}), a) | (g(\delta_t^{t+1}), a) \text{ is admissible w.r.t } W\}$. Any set that satisfies $W \subset B^{X^t}(W)$ is called *self-generating* at X^t. Furthermore recall that $\mathcal{V}(\delta_t^{t+1}) = \{V^{t,X^t}(f) | f \text{ is a subgame perfect equilibrium strategy profile}\}$.

Notice that

$$g(\delta_t^{t+1}) = (1-\hat{\delta}) \sum_{k=t+1}^{\infty} \mathrm{E}\left(\delta_t^k u\left(\pi^k(f)(X^k)\right) | \mathcal{F}_t\right),$$

for some strategy profile f. Furthermore since, δ_t^{t+1} is actually realized before the actions are taken and the multiplicative nature of our discount factor, the above equation becomes

$$g(\delta_t^{t+1}) = (1-\hat{\delta})\delta_t^{t+1}\left[u\left(\pi^{t+1}(f)(X^t)\right) + \sum_{k=t+2}^{\infty} \mathrm{E}\left(\delta_t^k u\left(\pi^k(f)(X^k)\right) | \mathcal{F}_t\right)\right]$$

which is equal to

$$g(\delta_t^{t+1}) = \delta_t^{t+1}\left[(1-\hat{\delta})u\left(\pi^{t+1}(f)(X^t)\right) + g(\delta_{t+1}^{t+2})\right]. \tag{8}$$

Now, it is easy to see that $\mathcal{V}(\delta_t^{t+1})$ is self-generating, as the pair, $(g(\delta_t^{t+1}), \pi^{t+1}(f)(X^t))$ is admissible with respect to $\mathcal{V}(\delta_t^{t+1})$ whenever f is a subgame perfect equilibrium strategy profile with

$$V^{t,X^t}(f) = (1-\hat{\delta})u(a_t) + g(\delta_t^{t+1}), \tag{9}$$

where $a_t = \pi^t(f)(X^{t-1})$. The two further points to notice is that, due to lemma 1 of Abreu, Pearce, and Stachetti (1990), $B^{X^t}(W)$ is compact whenever W is compact, and the operator B^{X^t} is monotone. Furthermore, since $\mathcal{V}(\delta_t^{t+1})$ is bounded (by $\mathcal{U}$), closure of $\mathcal{V}(\delta_t^{t+1})$, denoted by $\mathrm{cl}(\mathcal{V}(\delta_t^{t+1}))$ is compact. Hence, $B^{X^t}(\mathrm{cl}(\mathcal{V}(\delta_t^{t+1})))$ is compact, and due to

$\mathrm{cl}(\mathcal{V}(\delta_t^{t+1})) \subset B^{X^t}(\mathrm{cl}(\mathcal{V}(\delta_t^{t+1})))$ and self-generation $\mathrm{cl}(\mathcal{V}(\delta_t^{t+1})) \subset \mathcal{V}(\delta_t^{t+1})$. Thus, by Theorem 2 of Abreu, Pearce, and Stachetti (1990), $B^{X^t}(\mathcal{V}(\delta_t^{t+1})) = \mathcal{V}(\delta_t^{t+1})$, thus $\mathcal{V}(\delta_t^{t+1})$ is compact.

Lemma 2 *Suppose Assumptions 1, 2, 3, 4, 5 hold. Then, there exists $\underline{\delta} \in (0,1)$ such that for all $\delta_t^{t+1} \leq \underline{\delta}$ every subgame perfect strategy profile $f \in F$ is such that $f(X^t, a^t) \in A$ is a Nash equilibrium of G for any $t \in \mathbb{N}_0$.*

Proof. Without loss of generality, assume that the subgame perfect strategy f is such that $\left(\max_{a_i \in A_i} u_i(a_i, \pi_{-i}^t(f)(X^t)) - u_i(\pi^t(f)(X^t))\right) > 0$ for some $t \in \mathbb{N}_0$ for some X^t and for some $i \in N$. Because otherwise, the strategy is resulting in a repetition of period Nash behavior. Then, by equation 6, for any such subgame perfect strategy f and i and t and X^t

$$\delta_t^{t+1}\left(V_i^{t+1,X^t} - \mathrm{E}\left(v_i(X^{t+1})|\mathcal{F}_t\right)\right) \geq (1-\hat{\delta})\left(\max_{a_i \in A_i} u_i(a_i, \pi_{-i}^t(f)(X^t)) - u_i(\pi^t(f)(X^t))\right).$$

Both the left and the right hand sides of this inequality are strictly positive. Yet, when the prescribed action is not a Nash equilibrium of G, then the left hand side converges to 0 when δ_t^{t+1} tends to 0, but the right hand side is constant. ■

3 Inevitability of Nash behavior

In this section, we wish to present the main result of this study:

Theorem 1 *Suppose Assumptions 1, 2, 3, 4, 5 hold. Then for every $K \in \mathbb{N}$ and for every $\hat{\delta} \in (0,1)$ with $\hat{\delta} = \psi(\nu(d))$, there exists T which is almost surely in $\mathbb{N}_0$ with $\Pr\left(\mathbf{d}_\tau^{\tau+1} < \underline{\delta}\right) = 1$ for every $\tau = T, \ldots, T+K$.*

Using Lemma 2, the above Theorem establishes that in any equilibrium strategy, the occurrence of any finite number of consecutive repetitions of the period Nash action profile, must almost surely happen within a finite time window. That is, any equilibrium strategy almost surely entails arbitrary long consecutive observations of the period Nash action profile. In other words, every equilibrium strategy almost surely involves a stage, i.e. the stochastic process governing the one–shot discount factor possesses a *stopping time*, after which long consecutive repetitions of the period Nash action profile must be observed.

Proof of Theorem 1. Let $K \in \mathbb{N}$, and $\hat{\delta} \in (0,1)$ where $\hat{\delta} = \psi(\nu(d))$.

Without loss of generality assume that $\mathbf{X}$ has a support bounded from above. Recall that this situation corresponds to higher shocks to be interpreted as "positive" with respect to the probability of the continuation of the game, because such shocks will result in an increase in the current discount factor.

Define $\sup_{\omega \in \Omega} X(\omega) = \bar{X}$ as the highest realization of $\mathbf{X}$ with Ω denoting its space of events.

Similarly, let $\bar{Y} \in \{Y \in S : \psi_i(Y) < \underline{\delta}\}$, where $\underline{\delta}$ is as given in Lemma 2. Such a $\bar{Y}$ exists due to the third part of Assumption 4. It is appropriate to remind the reader that higher values of Y correspond to higher values of the resulting one–period discount factor.

Now, for any $K \in \mathbb{N}$ consider the state Y^* such that

$$\bar{Y} = \nu\left(\nu\left(\ldots\left(\nu\left(\nu\left(Y^*, \bar{X}\right), \bar{X}\right), \bar{X}\right)\ldots\right), \bar{X}\right),$$

where the compounding operation is performed $K+1$ times. Due to the third part of Assumption 4, such a Y^* exists.

Finally, define the *ergodic* state $\hat{Y} \in E$ (recall that $E \subset S$ denotes the set of ergodic states of S) such that $\hat{Y} \in \{Y \in E : Y < Y^*\}$. This set is non-empty because of the fourth part of Assumption 3.

Following Karlin and Taylor (1975), we define the Markovian time

$$\zeta \equiv \min\{\tau \in \mathbb{N}_0 : Y_\tau \leq \hat{Y}\}.$$

Then by construction, for any $k \leq K$ it must be that $\Pr(\psi_i(\mathbf{Y}_{\zeta+k}) \geq \underline{\delta}) = 0$. Finally due to ergodicity of $\hat{Y}_t$, ζ is a stopping time, it will almost surely happen in a finite time period, i.e. $\Pr(\zeta < \infty) = 1$. Hence, ζ is in $\mathbb{N}_0$ almost surely with $\Pr\left(\mathbf{d}_{\zeta+k}^{\zeta+k+1} < \underline{\delta}\right) = 1$ for any $k \leq K$. Indeed, this also implies that ζ is in $\mathbb{N}_0$ almost surely with $\mathrm{E}_\tau\left(\delta_\tau^{\tau+1}\right) < \underline{\delta}$ for every $\tau = \zeta, \ldots, \zeta + K$. ■

4 The Folk Theorem

In this section, we will present a subgame perfect Folk Theorem:

Theorem 2 *Suppose Assumptions 1, 2, 3, 4, 5 hold, and either* $\dim(\mathcal{U}) = n$ *or* $n = 2$ *and* $\mathcal{U}^0 \neq \emptyset$. *Then, for all* $\varepsilon > 0$, *there exists* $\bar{\delta} \in (0,1)$ *and* $d \in \mathbb{R}$ *such that for all* $u \in \mathcal{U}^0$ *and* $\hat{\delta} = \psi(\bar{\nu}(d))$ *with* $\hat{\delta} \geq \bar{\delta}$, *there exists a subgame perfect strategy* f *of* $G(d)$ *such that* $\|U(f, \hat{\delta}) - u\| < \varepsilon$.

Proof. Under these assumptions and Lemma 1, due to Fudenberg and Maskin (1986) and Fudenberg and Maskin (1991) we know that for all $\varepsilon > 0$, there exists $\bar{\delta} \in (0,1)$ and $d \in \mathbb{R}$ such that for all $u \in \mathcal{U}^0$ and $\hat{\delta} \geq \bar{\delta}$ with $\hat{\delta} = \psi(\bar{\nu}(d))$, there exists a strategy f of $G(d)$ such that:

$$V_i^{t,X^t}(f) \geq (1-\hat{\delta}) \max_{a_i \in A_i} u_i(a_i, \pi_{-i}(f)(X^t)) + \delta_t^{t+1} \mathrm{E}\left(v_i(X^{t+1}) | \mathcal{F}_t\right) \tag{10}$$

when δ_t^{t+1} that is associated with X^t is greater or equal to $\bar{\delta}$ and because of equation 1 $\delta_0^1 = \hat{\delta} = \mathrm{E}\left(\delta_t^{t+1} | \mathcal{F}_0\right)$ for all $t \in \mathbb{N}_0$. Thus, $V_i^{0,X^0}(f) = U(f, \hat{\delta})$, and $\|V_i^{0,X^0}(f) - u\| < \varepsilon$.

Hence, the remaining step is to construct the subgame perfect strategy (for states in which X^t have been such that in some period s, $s \leq t$, $\delta_s^{s+1} < \bar{\delta}$).

Having established that the set of subgame perfect equilibrium payoffs is compact for any given $\delta \in (0,1)$, we may employ a modification of simple strategies: Following Abreu (1988),

$f \in F$ is *a modified simple strategy profile represented by* $n+2$ *paths* $(\pi^{(0)}, \pi^{(1)}, \ldots, \pi^{(n)}, \pi^{NEq})$ if f specifies: If the shocks have all been such that the resulting one–period discount factors were greater or equal to $\bar{\delta}$ then: (i) play $\pi^{(0)}$ until some player deviates singly from $\pi^{(0)}$; (ii) for any $j \in N$, play $\pi^{(j)}$ if the jth player deviates singly from $\pi^{(i)}$, $i = 0, 1, \ldots, n$, where $\pi^{(i)}$ is the ongoing previously specified path; (iii) continue with the ongoing specified path $\pi^{(i)}$, $i = 0, 1, \ldots, n$, if no deviations occur or if two or more players deviate simultaneously; If the shocks have ever been such that in one of the previous periods the resulting one–period discount factor was strictly less than $\bar{\delta}$, play π^{NEq} consisting of the repetition of a^*, the Nash equilibrium pure action profile of the stage game. A profile $(\pi^{(0)}, \pi^{(1)}, \ldots, \pi^{(n)}, \pi^{NEq})$ of $n + 2$ outcome paths is *subgame perfect* if the modified simple strategy represented by it is a subgame perfect equilibrium.

Because of equations 1 and 2, the modified simple strategy defined above is subgame perfect: For any history in which the shocks have all been such that the resulting one–period discount factors were greater or equal to $\bar{\delta}$, the above specified modified simple strategy is a Nash equilibrium due to condition 6; and, for any other history the strategy specified requires the play of a Nash equilibrium pure action profile for the rest of the game, hence trivially is a Nash equilibrium in such subgames. ■

5 An Example: The Prisoners' Dilemma

The prisoners' dilemma is described as a normal form game G with two players ($N = \{1, 2\}$), each of whom have two actions: $A_i = \{C, D\}$ for $i = 1, 2$. Players' payoff functions are described by the following table:

$1\backslash 2$	C	D
C	$1, 1$	$-c, b$
D	$b, -c$	$0, 0$

where $b, c > 0$, $b > 1$, and $b < c + 2$ making sure that the cooperative payoff $(1, 1)$ is Pareto optimal and can be obtained only by (C, C). We denote player i's payoff function by $u_i : A \to \mathbb{R}$, for $i = 1, 2$, and $A = A_1 \times A_2$. We let $\mathcal{U}$ be the set of individually rational payoffs, i.e. $\mathcal{U} = \{u \in \text{con}(u(A)) : u_i \geq 0\}$, and the set of strictly individually rational payoffs takes the form $\mathcal{U}^0 = \{u \in \text{con}(u(A)) : u_i \geq 0\}$.

We let $\mathbf{X}_0 = 0$, and $\mathbf{X}_t = \mathbf{X}$ for $t > 0$ be identically and independently distributed steps of a standard random walk, i.e. X_t for $t > 1$ is a random variable that takes values in $\{-1, 1\}$ with a probability given by $\Pr(X_t = -1) = \Pr(X_t = 1) = \frac{1}{2}$. We interpret $X_t = 1$ as a positive (similarly $X_t = -1$ as a negative) *shock* at time t. Consequently, $Y_t = \sum_{s=1}^{t} X_s$ identifies the cumulative shocks up to period t.

By the Axiom of Choice, there exists a strictly increasing and with full support $\psi : \mathbb{Z} \to (0, 1) \cap \mathbb{Q}$ where $\mathbb{Q}$ denotes the set of rational numbers, such that $\{\psi \circ \mathbf{Y}_t\}_t$, hence $\{\mathbf{d}_t^{t+1}\}_t$, is a martingale.

Notice that Assumptions 1 and 2 are trivially satisfied. Regarding Assumption 3, the Markov property clearly holds because the transition of $\{\mathbf{Y}_t\}_t$, given by $\mathbf{Y}_t = \mathbf{Y}_{t-1} + \mathbf{X}_t$, for $t \in \mathbb{N}$. This also shows that its transition is clearly governed by $\{\mathbf{X}_t\}_t$ in a monotone (strictly increasing) fashion. The third part is also satisfied, because the covariance matrix of $\{\mathbf{Y}_t\}_t$ is non-negative by Feller (1950), which also shows that the fourth part of this same Assumption is satisfied since the entire state space of $\{\mathbf{Y}_t\}_t$ is ergodic. All the parts of Assumption 4 are satisfied due to the construction described in the previous paragraph. Finally, we will keep Assumption 5, requiring that every player forms his expectation with the most up to date σ-algebra.

Because that Assumptions 1 - 5 are satisfied, the characterization of subgame perfection given in equations 6 and 7 takes the following simple form when attention is restricted to the prisoners' dilemma: A strategy f is subgame perfect if and only if for all $i = 1, 2$ and for all $t = 0, 1, 2, \ldots$ and for all X^t with $X_k \in \{-1, 1\}$ such that $0 \le k \le t$, we have

$$V_i^{t,X^t}(f) \ge (1 - \hat{\delta}) \max_{a_i \in C,D} u_i(a_i, \pi_{-i}^t(f)(X^t)). \tag{11}$$

Notice that by Lemma 2, we know that there exists $\underline{\delta} > 0$ such that for all X^t with the resulting $\delta_t^{t+1} < \underline{\delta}$, any subgame perfect strategy must recommend the play of (D, D) which results in a payoff of 0 to both of the players.

For what follows, we will concentrate on obtaining the cooperative payoff exactly under subgame perfection.

First, we wish to identify the critical level of the discount factor that supports the cooperative payoff in subgame perfection. Notice that, due to condition 1 and 2, this critical level of discount factor is different from the one in the constant-discounting case, namely, $\frac{b-1}{b}$. Indeed, in this situation, this critical discount factor, that we call $\delta^\star$, is less than or equal to $\frac{b-1}{b}$.

Now, fix $t \in \mathbb{N}$, and X^t with a resulting $\delta_t^{t+1} \in (0, 1)$. Then, $V_i^{t,X^t}(f)$ where f is the simple strategy sustaining the cooperative payoff in subgame perfection and X^t is so that the resulting δ_t^{t+1} is high enough

$$(1 - \hat{\delta}) \left(1 + \mathrm{E}_t\left(\delta_t^{t+1}\right) + \mathrm{E}_t\left(\delta_t^{t+1}\delta_{t+1}^{t+2}\right) + \mathrm{E}_t\left(\delta_t^{t+1}\delta_{t+1}^{t+2}\delta_{t+2}^{t+3}\right) + \ldots\right) u_1(C, C).$$

Because that for all $\tau \ge t$, $\mathrm{E}_t\left(\prod_{k=t}^{\tau-1} \delta_k^{k+1}\right)$ is continuous in δ_t^{t+1}, there exists $\delta^\star(t) \in (0, 1)$ such that

$$(1 - \hat{\delta}) \left(1 + \mathrm{E}_t\left(\delta_t^{t+1}\right) + \mathrm{E}_t\left(\delta_t^{t+1}\delta_{t+1}^{t+2}\right) + \mathrm{E}_t\left(\delta_t^{t+1}\delta_{t+1}^{t+2}\delta_{t+2}^{t+3}\right) + \ldots\right) u_1(C, C) - (1 - \hat{\delta})b = 0.$$

This follows from the following: When δ_t^{t+1} is sufficiently close to 1, the left hand side of the the above equation is strictly positive. Moreover, when δ_t^{t+1} is sufficiently close to 0, then the left hand side of this equation is strictly negative. Thus, because of the Intermediate Value Theorem, there exists $\delta^\star(t)$, such that the left hand side is equal to zero. Moreover, because that $\{\mathbf{Y}_t\}_t$ has independent and stationary increments we obtain $\mathrm{E}\left(\mathbf{d}_{t+1}^{\tau}|\mathcal{F}_s\right) = \mathrm{E}\left(\mathbf{d}_{t+2}^{\tau+1}|\mathcal{F}_s\right)$ for all t, τ, s with $\tau > t \geq s$, which implies $\delta^\star(t) = \delta^\star(\tau)$ for all t, τ with $t \leq \tau$. Hence, we let $\delta^\star = \delta^\star(t)$ for all t.

Next, we wish to define the simple strategy supporting the cooperative payoff explicitly: Let $\pi^{(0)}$ be given by the repetitions of (C, C), $\pi^{(P)}$ the repetitions of (D, D), and finally π^{NEq} be also the repetitions of (D, D). Then the simple strategy is: If the shocks have all been such that the resulting one–period discount factors were greater or equal to $\delta^\star$, (i) play (C, C) until some player deviates singly from $\pi^{(0)}$ and (ii) play $\pi^{(P)}$ if there was a single player deviation, and (iii) continue playing $\pi^{(j)}$, $j = 0, P$ if there were either no deviations or multi-player deviations; and if there is a period in which the resulting one–period discount factors were strictly less than $\delta^\star$ play π^{NEq}, repetitions of (D, D), for the rest of the game.

Due to equations 1 and 2, the above strategy f is such that $V_i^{0,X^0}(f) = 1$ for all i whenever $\delta_0^1 = \hat{\delta} \geq \delta^\star$. Moreover, it easily can be seen that f is subgame perfect.

But, by Lemma 2 and Theorem 1, we know that in any subgame perfect equilibrium strategy f, for every $K \in \mathbb{N}$ and for every $\hat{\delta} \in (0, 1)$ with $\hat{\delta} = \delta_0^1$, there exists T which is almost surely in $\mathbb{N}_0$ and the period behavior prescribed by f must be (D, D) for every $\tau = T, \ldots, T + K$.

References

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

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

Aumann, R., and L. Shapley (1994): "Long-Term Competition – A Game-Theoretic Analysis," in *Essays in Game Theory in Honor of Michael Maschler*, ed. by N. Megiddo. Springer-Verlag, New York.

Barlo, M., G. Carmona, and H. Sabourian (2007): "Bounded Memory with Finite Action Spaces," Sabancı University, Universidade Nova de Lisboa and University of Cambridge.

——— (2009): "Repeated Games with One – Memory," *Journal of Economic Theory*, 144, 312–336.

Baye, M., and D. W. Jansen (1996): "Repeated Games with Stochastic Discounting," *Economica*, 63(252), 531–541.

Dutta, P. (1995): "A Folk Theorem for Stochastic Games," *Journal of Economic Theory*, 66, 1–32.

Feller, W. (1950): *An Introduction to Probability Theory and Its Applications Volume I.* John Wiley and Sons, 3^{rd} edn.

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

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

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

Fudenberg, D., and Y. Yamamato (2010): "The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring," Harvard University.

Hansen, L., and S. Richard (1987): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55, 587–614.

Harrison, J. M., and D. Kreps (1979): "Martingale and Arbitrage in Multi-Period Securities Markets," *Journal of Economic Theory*, 20, 381–408.

Hörner, J., and W. Olszewski (2006): "The Folk Theorem for Games with Private Almost-Perfect Monitoring," *Econometrica*, 74, 1499–1544.

Hörner, J., T. Sugaya, S. Takahashi, and N. Vieille (2010): "Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \to 1$ and a Folk Theorem," Yale University, Princeton University, Princeton University, and HEC.

Kalai, E., and W. Stanford (1988): "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica*, 56, 397–410.

Karlin, S., and H. M. Taylor (1975): *A First Course in Stochastic Processes*. Academic Press, 2 edn.

Mailath, G., and W. Olszewski (2008): "Folk Theorems with Bounded Recall under (Almost) Perfect Monitoring," University of Pennsylvania and Northwestern University.

Osborne, M., and A. Rubinstein (1994): *A Course in Game Theory*. MIT Press, Cambridge.

Ross, S. A. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1), 97–109.

SABOUORIAN, H. (1998): "Repeated Games with M-period Bounded Memory (Pure Strategies)," *Journal of Mathematical Economics*, 30, 1–35.

SUGAYA, T. (2010): "Characterizing the Limit Set of PPE Payoffs with Unequal Discounting," Princeton University.

WILLAMS, D. (1991): *Probability with Martingales*. Cambridge University Press.