Automatica 183 (2026) 112656

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Stochastic games of risk-sensitive players with quasi-hyperbolic discounting ☆

Anna Jaśkiewicz [a], Andrzej S. Nowak [b], Łukasz Woźny [c,*]

[a] Faculty of Pure and Applied Mathematics, Wrocław University of Science and Technology, Wrocław, Poland
[b] Institute of Mathematics, University of Zielona Góra, Zielona Góra, Poland
[c] Department of Quantitative Economics, SGH Warsaw School of Economics, Poland

ARTICLE INFO

Article history:
Received 11 September 2024
Received in revised form 22 May 2025
Accepted 11 September 2025

Keywords:
Quasi-hyperbolic discounting
Stochastic game
Risk-sensitive player
Markov perfect equilibrium

ABSTRACT

In this paper, we study a stochastic game on a countable state space with quasi-hyperbolic discounting. In each period, there are N selves who play a non-cooperative game. It is assumed that an individual player, represented by a sequence of autonomous selves, is equipped with a risk-sensitive coefficient and evaluates their expected payoff using the exponential certainty equivalent. The main objective is to prove the existence of a Markov perfect equilibrium in an infinite horizon model. If the transition probabilities, one-stage payoff functions, and action sets for the selves do not depend on the period number, then the game possesses a stationary Markov perfect equilibrium. A dynamic game on networks, belonging to the class of stochastic games studied in this paper, is also discussed.

1. Introduction

Empirical and experimental studies on time preferences have found that discount rates are much higher in the short run than in the long run, leading to time-inconsistent preferences. These preferences have been observed in numerous cases and individual behaviors; see, for instance, Krusell, Kuruşçu, and Smith (2002), Krusell and Smith (2003), Phelps and Pollak (1968), and Strotz (1956). To model this phenomenon, researchers have adopted discount functions from the class of generalized hyperbolas. The discrete-time analog of quasi-hyperbolic discounting involves the coefficients:

$$1, \alpha\beta, \alpha\beta^2, \dots,$$

where $\beta \in (0,1)$ is a long-run discount factor, and $\alpha > 0$ is a short-run discount factor. Such discounting was first used by Phelps and Pollak (1968) in the context of optimal economic growth with deterministic transition functions. They noted that a time-consistent solution could be achieved by finding an equilibrium in certain games played by countably many short-lived players, each of whom acts only once. Furthermore, Harris and Laibson (2001) studied a consumption-saving model under uncertainty with "hyperbolic consumers". The state space in their model consists of positive real numbers, and the transition is described by a difference equation with i.i.d. non-atomic shocks having a distribution with bounded support. A non-stationary consumption-saving model with more general non-atomic transition laws was examined by Balbus, Jaśkiewicz, and Nowak (2020), who demonstrated the existence of a Markov perfect equilibrium. A survey of many related results and a comprehensive discussion on time-consistent solutions in dynamic decision models can be found in Balbus, Reffett, and Woźny (2023). Additionally, Alj and Haurie (1983) extended the approach of Phelps and Pollak (1968) to characterize intergenerational equilibria in dynamic stochastic games with finite state and action spaces.

Dynamic inconsistency is now playing an increasingly important role in many fields. It has inspired algorithms based on reinforcement learning for the computation of Markov perfect equilibria; see, for example, the recent contribution by Eshwar, Motwani, Roy, and Thoppe (2024). Moreover, Haurie (2005) and Haurie and Moresino (2006) have explored compelling applications of time inconsistency in the context of global climate change problems. Readers are also referred to other works, such as Björk, Khapko, and Murgoci (2021), Christensen and Lindensjö (2018), which examine various related control problems for models with a general state space.

Empirical evidence suggests that individuals tend to be risk-averse or are even compelled to be so by regulations, particularly in sectors such as finance and insurance. The seminal papers by Howard and Matheson (1972), Jacobson (1973), and Jaquette (1973, 1976) marked the beginning of extensive research on dynamic programming with risk-sensitive preferences of decision makers. Loosely speaking, risk sensitivity weights the possible

☆ The work is supported by the National Science Centre, Poland Grant UMO-2022/47/B/HS4/00331. The material in this paper was not presented at any conference This paper was recommended for publication in revised form by Associate Editor Dena Firoozi under the direction of Editor Florian Dorfler.

* Corresponding author.
E-mail addresses: anna.jaskiewicz@pwr.edu.pl (A. Jaśkiewicz), a.nowak@im.uz.zgora.pl (A.S. Nowak), lukasz.wozny@sgh.waw.pl (Łu. Woźny).

https://doi.org/10.1016/j.automatica.2025.112656
0005-1098/

fluctuations around the mean. A straightforward approach to modeling this involves using a weighted criterion that includes both the expectation and the variance of a random income. Today, the term risk-sensitive decision maker primarily refers to settings in which the expected value of a random payoff is replaced by the certainty equivalent derived from an exponential utility function. In this context, all moments of the random payoff are taken into account. This is evident from the Taylor series expansion of the exponential and logarithmic functions around zero. The negative of the certainty equivalent for an exponential utility function is also known as the *entropic risk measure* (Föllmer & Schied, 2004).

In the past decade, there has been significant progress in considering the risk-averse attitudes of decision makers. For a comprehensive overview, readers are referred to the recent survey by Bäuerle and Jaśkiewicz (2024), which includes an extensive and up-to-date list of references. The approach based on exponential utility has proven useful in various fields, including economics (Jaśkiewicz & Nowak, 2014; Sargent & Stachurski, 2023), actuarial science (Bäuerle & Jaśkiewicz, 2015), finance (Föllmer & Schied, 2004; Pitera & Stettner, 2023), and operations research (Hansen & Sargent, 1995; Jacobson, 1973; Whittle, 1990). Moreover, as argued by Hansen and Sargent (1995) and Başar (2021), risk-sensitive preferences are particularly appealing because they can also model preferences for robustness. Within this framework, the parameter in the exponential utility function is interpreted as a robustness parameter. These preferences are not only practically relevant but also mathematically rich, inspiring numerous studies in the context of Markov decision processes (Bäuerle & Rieder, 2014; Chung & Sobel, 1987; Di Masi & Stettner, 2000) and dynamic games (Başar, 1999; Basu & Ghosh, 2018; Caravani & Papavassilopoulos, 1990; Klompstra, 2000; Solano, 2014; Solano & Shevkoplyas, 2011).

This paper introduces a new class of stochastic games on a countable state space involving risk-averse players who employ quasi-hyperbolic discounting. The one-stage payoffs and transition probabilities may vary over time. The primary objective is to establish the existence of a Markov perfect equilibrium (MPE). Our proof approach is inspired by value function iteration in discounted dynamic programming for finite horizon games. We construct a sequence of finite-step equilibria and demonstrate that a convergent subsequence yields an equilibrium as the time horizon tends to infinity. For the stationary case, we establish the existence of a stationary MPE using a fixed-point theorem. Our work generalizes the model studied in Basu and Ghosh (2018), where standard discounted stochastic games with risk-sensitive players are analyzed. Basu and Ghosh (2018) proved the existence of Nash equilibria within the class of Markov strategies. Additionally, our paper draws significantly from the work of Alj and Haurie (1983), who investigated MPE in stochastic game models with risk-neutral players across generations. Specifically, they proved that every finite horizon game admits an equilibrium in Markov strategies, while an infinite horizon game with stationary data admits a stationary MPE. Their finite horizon result is based on a dynamic programming algorithm, whereas the infinite horizon result relies on a fixed-point argument. However, the asymptotic behavior of equilibria and the value (or equilibrium) functions as the time horizon tends to infinity was not addressed in Alj and Haurie (1983). In this paper, we demonstrate that accumulation points of finite horizon equilibria — even in non-stationary models — constitute MPE in infinite horizon risk-sensitive games. Moreover, since we deal with expected multiplicative payoffs, we require continuity lemmas for the expected payoffs in stationary strategies. These strategies are endowed with a natural compact topology, which is not used in the risk-neutral setting. Finally, we note that MPEs in deterministic (pure) strategies for risk-sensitive Markov decision processes with quasi-hyperbolic discounting were studied in Jaśkiewicz and Nowak (2024).

The paper is organized as follows. Basic notation is introduced in Section 2. Section 3 presents the model along with the underlying assumptions and the formulation of the main results. The proofs are provided in Section 4. Finally, Section 5 illustrates an application of our model to a stochastic network formation process. To the best of our knowledge, no existing literature addresses the existence of stationary Markov perfect equilibria in stochastic network formation games involving risk-sensitive and potentially quasi-hyperbolic players. Our results are also applicable to a class of stochastic games modeling imperfect competition between firms — such as those discussed in Escobar (2013) — where firms are managed by risk-sensitive agents who evaluate future profit streams using quasi-hyperbolic discounting.

2. Preliminaries and notation

We use $\mathbb{R}$, $\mathbb{R}_+$ and $\mathbb{N}$ to denote the sets of all real numbers, positive real numbers and positive integers, respectively. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let Z be a bounded random *payoff* of a decision maker defined on Ω. A *certainty equivalent* is a concept in economics that refers to the guaranteed amount of money or value, an individual would consider equally desirable or preferable to a risky or uncertain return from a future inflow. It helps to evaluate and quantify the value of uncertainty and enables the individual to make decisions based on their own risk preferences. Assume that an individual has an exponential utility function $\mathcal{U}(z) = re^{rz}$ with a fixed coefficient $r \in \mathbb{R}\setminus\{0\}$. A *certainty equivalent* of Z is the number $c_e(r, Z)$ such that $\mathcal{U}(c_e(r, Z)) = \mathbb{E}\mathcal{U}(Z)$. Observe that

$$c_e(r, Z) = \frac{1}{r} \ln \mathbb{E}(e^{rZ}).$$

The negative of $c_e(r, Z)$ is also known as the entropic risk measure of Z, see, e.g., Föllmer and Schied (2004). The parameter $-r$ is called the Arrow–Pratt risk coefficient (see Föllmer and Schied (2004) and Pratt (1964)). An individual equipped with risk sensitivity coefficient $-r > 0$ ($r > 0$) is risk-averse (risk-loving). Basically, this parameter shows the degree of risk sensitivity of a decision maker. Note that by applying the Taylor expansion around $r = 0$ for $\mathcal{U}$ we obtain

$$c_e(r, Z) \approx \mathbb{E}Z + \frac{r}{2} VarZ.$$

Therefore, if $-r > 0$, then the individual who considers $c_e(r, Z)$ thinks not only of the expected value $\mathbb{E}Z$ of the random utility (payoff) Z, but also of its variance.

Let K be any compact metric space. By $C(K)$ we denote the Banach space of all continuous real-valued functions on K endowed with the maximum norm. By the Riesz representation theorem (Corollary 14.15 in Aliprantis and Border (2006)) the topological dual space $C^*(K)$ of $C(K)$ consists of all finite countable additive regular measures on the Borel sets in K. The space $C^*(K)$ is endowed with the weak-star metrizable topology. It is then a locally convex linear topological space. The set $\Pr(K)$ of all probability measures on K is a compact convex subset of $C^*(K)$ (Theorem 6.21 in Aliprantis and Border (2006)). We recall that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures in $\Pr(K)$ converges in the weak-star sense to μ, which we denote as $\mu_n \to^* \mu$, if $\lim_{n\to\infty} \int_K f d\mu_n = \int_K f d\mu$ for each $f \in C(K)$, see Proposition 7.21 in Bertsekas and Shreve (1978).

If K is a metric space, then by $B(K)$ we denote the Banach space of all bounded real-valued functions on K endowed with the supremum norm $\|f\| = \sup_{y\in K} |f(y)|$.

3. The model and main results

In the *N-person stochastic game model with quasi-hyperbolic discounting*, we envision each individual player $i \in \mathcal{N} := \{1, 2, \ldots, N\}$ as a sequence $(i_t)_{t\in\mathbb{N}}$ of autonomous temporal *selves*. For each $t \in \mathbb{N}$ we denote by $G_t := \{1_t, 2_t, \ldots, N_t\}$ the set of N selves who act in tth period of the game. Usually, a self $i_t \in G_t$ is also called a player i_t. Then, the set G_t can be treated as the set of players. The subscript t indicates that they live and choose actions in period t.

We assume that the players play a discrete-time non-cooperative stochastic game *a là* Shapley (1953) in a framework similar to that of Alj and Haurie (1983).[1] A new feature in our approach is that we assume that the players are risk-averse, which is modeled by using an exponential certainty equivalent. Moreover, we allow the one-stage payoff functions and transition probabilities to depend on the time period.

The *state space* X in the game is countable. For every $i_t \in G_t$, *action space* A_{i_t} is compact metric and is endowed with its Borel σ-algebra. The set $A_{i_t}(x)$ is a non-empty *compact* subset of A_{i_t}, $x \in X$, $i_t \in G_t$. We define

$$A_t := \prod_{i=1}^{N} A_{i_t} \quad \text{and} \quad A_t(x) := \prod_{i=1}^{N} A_{i_t}(x).$$

Then $A_{i_t}(x)$ is the set of all *feasible actions* for player $i_t \in G_t$ in state $x \in X$. We also define

$$\mathbb{K}_t = \{(x, \boldsymbol{a}_t) : x \in X,\ \boldsymbol{a}_t = (a_{1_t}, \ldots, a_{N_t}) \in A_t(x)\}.$$

The *one-stage payoff function* for self i_t (in period t) is $u_{i_t} : \mathbb{K}_t \to \mathbb{R}$. The *transition probability* from x to $y \in X$, when the players choose a profile $\boldsymbol{a}_t = (a_{1_t}, a_{2_t}, \ldots, a_{N_t})$ of actions in $A_t(x)$ is denoted by $q_t(y|x, \boldsymbol{a}_t)$.

We impose the following additional assumptions.

Assumption A. **(i)** There exists $c > 0$ such that

$$\|u_{i_t}\| = \sup_{(x,\boldsymbol{a}_t)\in\mathbb{K}_t} |u_{i_t}(x, \boldsymbol{a}_t)| \le c$$

for all i_t, $t \in \mathbb{N}$ and $i \in \mathcal{N}$. Moreover, the functions $u_{i_t}(x, \cdot)$ are continuous on $A_t(x)$ for every $x \in X$.

(ii) The function $q_t(y|x, \cdot)$ is continuous on $A_t(x)$ for all $x, y \in X$.

In this paper we establish the existence of a Markov perfect equilibrium in the game under consideration. Therefore, we restrict attention to Markov strategies of the players. Let Φ_{i_t} be the set of all transition probabilities φ_{i_t} from X to A_{i_t} such that $\varphi_{i_t}(A_{i_t}(x)|x) = 1$ for each $x \in X$. An element $\varphi_{i_t} \in \Phi_{i_t}$ is a *Markov strategy* for self (or player) $i_t \in G_t$. We put

$$\Phi_t := \prod_{i=1}^{N} \Phi_{i_t}. \tag{1}$$

An element of Φ_t is denoted by $\boldsymbol{\varphi}_t = (\varphi_{1_t}, \ldots, \varphi_{N_t})$ and is called a *Markov strategy profile* of the players (selves) in G_t. A *Markov strategy* for player i is represented by a sequence $(\varphi_{i_t})_{t\in\mathbb{N}}$ of Markov strategies for selves $(i_t)_{t\in\mathbb{N}}$. Here $\varphi_{i_t} \in \Phi_{i_t}$ for each $t \in \mathbb{N}$. Then, $(\varphi_{i_\tau})_{\tau\ge t+1}$ is a Markov strategy for all selves $(i_\tau)_{\tau\ge t+1}$ following player i_t. Put $\Phi_{[t,\infty)} := \prod_{\tau=t}^{\infty} \Phi_\tau$. Next define the sequence $\boldsymbol{\varphi}_{[t,\infty)} := (\boldsymbol{\varphi}_t, \boldsymbol{\varphi}_{t+1}, \ldots) \in \Phi_{[t,\infty)}$ of Markov strategy profiles used by players from G_τ with $\tau \ge t$.

Assume that $\boldsymbol{\varphi}_t = (\varphi_{1_t}, \varphi_{2_t}, \ldots, \varphi_{N_t}) \in \Phi_t$. Then

$$\boldsymbol{\varphi}_t(x) := (\varphi_{1_t}(\cdot|x), \varphi_{2_t}(\cdot|x), \ldots, \varphi_{N_t}(\cdot|x))$$

with $\varphi_{i_t}(\cdot|x) \in \Pr(A_{i_t}(x))$, $i \in \mathcal{N}$, is the profile of mixed strategies used by players from G_t in state $x \in X$, and

$$\boldsymbol{\varphi}_t(d\boldsymbol{a}_t|x) := \varphi_{1_t}(da_{1_t}|x) \otimes \varphi_{2_t}(da_{2_t}|x) \otimes \cdots \otimes \varphi_{N_t}(da_{N_t}|x)$$

is the product probability measure on $A_t(x)$ induced by $\varphi_{i_t}(da_{i_t}|x)$, $i_t \in G_t$.

Let $H_{[t,\infty)} = \prod_{\tau=t}^{\infty} \mathbb{K}_\tau$ be the space of all infinite histories of the game (plays) endowed with the product σ-algebra. For any $\boldsymbol{\varphi}_{[1,\infty)} \in \Phi_{[1,\infty)}$ and state $x_t \in X$ in period $t \in \mathbb{N}$, a probability measure $\mathbb{P}_{x_t}^{\boldsymbol{\varphi}_{[t,\infty)}}$ and a stochastic process $(x_\tau, \boldsymbol{a}_\tau)_{\tau\ge t}$ are defined on $H_{[t,\infty)}$ in a canonical way according to the Ionescu–Tulcea theorem (see Proposition 7.28 in Bertsekas and Shreve (1978)). The expectation operator with respect to $\mathbb{P}_{x_t}^{\boldsymbol{\varphi}_{[t,\infty)}}$ is denoted by $\mathbb{E}_{x_t}^{\boldsymbol{\varphi}_{[t,\infty)}}$.

The *risk-sensitive total payoff (utility)* of player $i_t \in G_t$ depends on his current one-stage payoff and one-stage payoffs to be received by all following selves $(i_\tau)_{\tau\ge t+1}$. It is of the form

$$U_{i_t}(\boldsymbol{\varphi}_{[t,\infty)})(x_t) := U_{i_t}^r(\boldsymbol{\varphi}_{[t,\infty)})(x_t) := \tag{2}$$
$$\frac{1}{r} \ln \mathbb{E}_{x_t}^{\boldsymbol{\varphi}_{[t,\infty)}} \exp\left(r u_{i_t}(x_t, \boldsymbol{a}_t) + \alpha\beta \sum_{\tau=t+1}^{\infty} \beta^{\tau-t-1} r u_{i_\tau}(x_\tau, \boldsymbol{a}_\tau) \right),$$

where $\beta \in (0, 1)$ is called a *long-run discount factor*, $\alpha \in (0, 1]$ is called a *short-run discount factor*, and $r < 0$ represents *risk coefficient* of the players. Note that the risk aversion increases when r decreases, i.e., $r \to -\infty$.

The idea of using quasi-hyperbolic discounting in the risk-neutral setting, i.e., without an exponential certainty equivalent as in (2), goes back among others to Alj and Haurie (1983), Harris and Laibson (2001), Jaśkiewicz and Nowak (2021), Phelps and Pollak (1968). More precisely, by the risk-neutral case we mean the model, where $U_{i_t}(\boldsymbol{\varphi}_{[t,\infty)})(x_t)$ is replaced by

$$U_{i_t}^0(\boldsymbol{\varphi}_{[t,\infty)})(x_t) :=$$
$$\mathbb{E}_{x_t}^{\boldsymbol{\varphi}_{[t,\infty)}} \left(u_{i_t}(x_t, \boldsymbol{a}_t) + \alpha\beta \sum_{\tau=t+1}^{\infty} \beta^{\tau-t-1} u_{i_\tau}(x_\tau, \boldsymbol{a}_\tau) \right). \tag{3}$$

A related work on multi-generational games of resource extraction with a continuum of states, specific transition probabilities and the expected payoff criterion as in (3) was studied in Balbus and Nowak (2008).

It is well-known that

$$U_{i_t}^0(\boldsymbol{\varphi}_{[t,\infty)})(x_t) = \lim_{r\to 0^-} U_{i_t}^r(\boldsymbol{\varphi}_{[t,\infty)})(x_t). \tag{4}$$

Let $(\boldsymbol{\varphi}_t)_{t\in\mathbb{N}} \in \Phi_{[1,\infty)}$ and $k > t+1$. For $\boldsymbol{\varphi}_{[t+1,k]} := (\boldsymbol{\varphi}_{t+1}, \ldots, \boldsymbol{\varphi}_k)$ define

$$V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,k]})(x_{t+1}) :=$$
$$\mathbb{E}_{x_{t+1}}^{\boldsymbol{\varphi}_{[t+1,k]}} \exp\left(\alpha \sum_{\tau=t+1}^{k} \beta^{\tau-t} r u_{i_\tau}(x_\tau, \boldsymbol{a}_\tau) \right). \tag{5}$$

Moreover, let

$$V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)})(x_{t+1}) :=$$
$$\mathbb{E}_{x_{t+1}}^{\boldsymbol{\varphi}_{[t+1,\infty)}} \exp\left(\alpha \sum_{\tau=t+1}^{\infty} \beta^{\tau-t} r u_{i_\tau}(x_\tau, \boldsymbol{a}_\tau) \right). \tag{6}$$

Note that the right-hand side in (6) depends on strategy profiles and one-stage payoffs of all selves $(i_\tau)_{\tau\ge t+1}$ following self i_t. The analogous remark concerns the formula in (5). Then

$$U_{i_t}(\boldsymbol{\varphi}_{[t,\infty)})(x_t) = \frac{1}{r} \ln \int_{A_t(x_t)} \exp(r u_{i_t}(x_t, \boldsymbol{a}_t)) \times$$

[1] One should note that Alj and Haurie (1983) studied a multi-generational stochastic game. However, their results can also be re-formulated for a stochastic game with quasi-hyperbolic discounting.

$$\sum_{x_{t+1}\in X} V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)})(x_{t+1})q_t(x_{t+1}|x_t,\boldsymbol{a}_t)\boldsymbol{\varphi}_t(d\boldsymbol{a}_t|x_t).$$

If $\boldsymbol{\varphi}_t = (\varphi_{1_t}, \varphi_{2_t}, \ldots, \varphi_{N_t}) \in \Phi_t$ and $\xi_{i_t} \in \Phi_{i_t}$, then $[\boldsymbol{\varphi}_{-i_t}, \xi_{i_t}]$ is the strategy profile $\boldsymbol{\varphi}_t$ of players from G_t with φ_{i_t} replaced by ξ_{i_t}. Moreover, $[\boldsymbol{\varphi}_{-i_t}, \xi_{i_t}](x) := (\phi_{1_t}(\cdot|x), \phi_{2_t}(\cdot|x), \ldots, \phi_{N_t}(\cdot|x))$ with $\phi_{j_t}(\cdot|x) = \varphi_{j_t}(\cdot|x)$ for $j_t \neq i_t$ and $\phi_{i_t}(\cdot|x) = \xi_{i_t}(\cdot|x)$. In addition, we also put

$$[\boldsymbol{\varphi}_{-i_t}, \xi_{i_t}](d\boldsymbol{a}_t|x) :=$$
$$\phi_{1_t}(da_{1_t}|x) \otimes \phi_{2_t}(da_{2_t}|x) \otimes \cdots \otimes \phi_{N_t}(da_{N_t}|x)$$

with ϕ_{i_t} defined above. This is the product measure on $A_t(x)$ induced by the strategy profile $[\boldsymbol{\varphi}_{-i_t}, \xi_{i_t}]$.

Definition 1. A sequence $(\boldsymbol{\varphi}_t^*)_{t\in\mathbb{N}} \in \Phi_{[1,\infty)}$ constitutes a *Markov perfect equilibrium* (*MPE*) in the stochastic game with quasi-hyperbolic discounting, if for each $t \in \mathbb{N}$, every player $i_t \in G_t$ and a Markov strategy $\xi_{i_t} \in \Phi_{i_t}$ it holds

$$U_{i_t}(\boldsymbol{\varphi}^*_{[t,\infty)})(x_t) \geq U_{i_t}([\boldsymbol{\varphi}^*_{-i_t}, \xi_{i_t}], \boldsymbol{\varphi}^*_{[t+1,\infty)})(x_t), \tag{7}$$

for all $x_t \in X$.

Before stating our first main result, we give an equivalent formulation of the above definition. Let $(\boldsymbol{\varphi}_t)_{t\in\mathbb{N}} \in \Phi_{[1,\infty)}$, $x_t \in X$ and $t \in \mathbb{N}$. With these data we associate an N-person one-shot game $\Gamma(x_t, \boldsymbol{\varphi}_{[t+1,\infty)})$ where the payoff function to player $i_t \in G_t$ under the strategy profile $\boldsymbol{a}_t = (a_{1_t}, a_{2_t}, \ldots, a_{N_t}) \in A_t(x_t)$ is

$$P_{i_t}(x_t, \boldsymbol{\varphi}_{[t+1,\infty)}, \boldsymbol{a}_t) := \tag{8}$$
$$-\exp(ru_{i_t}(x_t,\boldsymbol{a}_t)) \sum_{x_{t+1}\in X} V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)})(x_{t+1})q_t(x_{t+1}|x_t,\boldsymbol{a}_t).$$

Since $r < 0$ and the logarithm function $\ln(\cdot)$ is increasing, we obtain the following equivalent definition.

Lemma 1. *A sequence $\boldsymbol{\varphi}^*_{[1,\infty)} = (\boldsymbol{\varphi}_t^*)_{t\in\mathbb{N}} \in \Phi$ is an MPE in the stochastic game with quasi-hyperbolic discounting, if and only if for each $t \in \mathbb{N}$ and $x_t \in X$, $\boldsymbol{\varphi}_t^*(x_t)$ is a Nash equilibrium in the one-shot game $\Gamma(x_t, \boldsymbol{\varphi}^*_{[t+1,\infty)})$.*

Here is our first main result.

Theorem 1. *Under Assumption A the stochastic game with quasi-hyperbolic discounting has an MPE.*

In a stationary model of a stochastic game with quasi-hyperbolic discounting we have $A_{i_t} = A_i$, $A_{i_t}(x) = A_i(x)$ for all $t \in \mathbb{N}$ and $x \in X$. Thus, we can write $\mathbb{K}_t = \mathbb{K}$. By Φ_i we denote the set of all transition probabilities ϕ_i from X to A_i such that $\phi_i(A_i(x)|x) = 1$ for all $x \in X$. Therefore, $\Phi_{i_t} = \Phi_i$ for every $t \in \mathbb{N}$. Moreover, we also have $u_{i_t} = u_i$, and $q_t = q$ for all $t \in \mathbb{N}$.

Definition 2 (*SMPE*)**.** A *stationary Markov perfect equilibrium* is *an MPE $\boldsymbol{\varphi}^* = (\boldsymbol{\varphi}_t^*)_{t\in\mathbb{N}}$ such that $\boldsymbol{\varphi}_t^* = \boldsymbol{\varphi}_1^*$ for all $t \in \mathbb{N}$.*

Theorem 2. *Under Assumption A made for a stationary stochastic game with quasi-hyperbolic discounting, there exists an SMPE.*

Remark 1. The proof of Theorem 1 consists of two steps. First by backward algorithm we obtain *MPE* for the finite time horizon games. This result extends Theorem 4.1 in Alj and Haurie (1983) stated for finite Markov games within a risk-neutral framework. Then, an *MPE* in Theorem 1 is obtained as an accumulation point of equilibria in finite time horizon games. Theorem 2, on the other hand, is based on a fixed point argument and extends Theorem 5.1 in Alj and Haurie (1983) proved for the risk-neutral case.

Remark 2. For convenience of notation we assume that discount factors and risk coefficients are the same for all players. However, our results remain true when we allow them to depend on $i \in \mathcal{N}$.

We close this section with some remarks on stationary model with standard discounting, i.e., $\alpha = 1$ in (2). We have $G_t = \mathcal{N}$ for all $t \in \mathbb{N}$. A *Markov strategy* for player $i \in \mathcal{N}$ is a sequence $(\phi_t)_{t\in\mathbb{N}}$ where $\phi_t \in \Phi_i$ for all $t \in \mathbb{N}$.

Remark 3. The existence of Nash equilibrium in Markov strategies in risk-sensitive games with standard discounting was established in Basu and Ghosh (2018). Their proof combines the corresponding Bellman equations established for each player with the Kakutani–Fan–Glicksberg fixed point theorem applied to an appropriately defined best-response correspondence. A stationary Nash equilibrium in the model of Basu and Ghosh (2018) fails to exist. Even for one-person game (a Markov decision process) with finite state and action spaces one can only show the existence of a Markov optimal policy (strategy). Indeed, the example in Jaquette (1976) with a small correction given in Bäuerle and Jaśkiewicz (2024) illustrates this fact. Our Theorem 2 shows that a stationary *MPE* exists (also for $\alpha = 1$) if we accept a solution concept from Definition 2. It is weaker than a Nash equilibrium obtained in Basu and Ghosh (2018), however, it is time-consistent. Note that in case of an *MPE*, the player in generation t may deviate from the assigned strategy only in period t. In case of a Nash equilibrium, the player may change his strategy at any time period.

4. Proofs

We start with some auxiliary facts and notation. Assume that $t \in \mathbb{N}$ and $i_t \in G_t$. Let $C^*_{i_t} := \prod_{x\in X} C^*(A_{i_t}(x))$ be the countable product of topological dual spaces $C^*(A_{i_t}(x))$ of the Banach spaces $C(A_{i_t}(x))$. Assume that $C^*_{i_t}$ is endowed with the product of weak-star topologies on the spaces $C^*(A_{i_t}(x))$. Since all action spaces $A_{i_t}(x)$ are compact metric, so $C^*_{i_t}$ is metrizable. It is also a locally convex linear topological space, see Chapter 5.14 in Aliprantis and Border (2006). Note that by the Riesz representation theorem (Corollary 14.15 in Aliprantis and Border (2006)), every $\varphi_{i_t} = (\varphi_{i_t}(\cdot|x))_{x\in X} \in \Phi_{i_t}$ can be recognized as an element of a sequentially compact convex set $\prod_{x\in X} \Pr(A_{i_t}(x)) \subset C^*_{i_t}$. Sequential compactness of the product space follows from Tychonoff's theorem. Therefore, we can think that Φ_{i_t} is a compact convex subset of a locally convex metrizable vector space. The convergence $\varphi^n_{i_t} \to \varphi^o_{i_t}$ in Φ_{i_t} means that $\varphi^n_{i_t}(\cdot|x) \to^* \varphi^o_{i_t}(\cdot|x)$ as $n \to \infty$ for all $x \in X$. Assume that Φ_t defined in (1) is given the product topology. Let $\boldsymbol{\varphi}^n_t \to \boldsymbol{\varphi}^o_t$ in Φ_t as $n \to \infty$ and $g \in C(A_{t_1}(x) \times \cdots \times A_{t_N}(x))$ for every $x \in X$. By Proposition 7.21 and Lemma 7.12 in Bertsekas and Shreve (1978), we have

$$\lim_{n\to\infty} \int_{A_t(x)} g(\boldsymbol{a})\boldsymbol{\varphi}^n_t(d\boldsymbol{a}|x) = \int_{A_t(x)} g(\boldsymbol{a})\boldsymbol{\varphi}^o_t(d\boldsymbol{a}|x), \tag{9}$$

for all $x \in X$. Moreover, let us endow the sets $\Phi_{[t,\infty)}$ and $\Phi_{[t,T]}$ for any $T \geq t$ with the product topology. Then, by Tychonoff's theorem $\Phi_{[t,\infty)}$ and $\Phi_{[t,T]}$ are sequentially compact as well.

Lemma 2. *Let Assumption A hold. Assume that $v^n, v \in B(X)$ and $\lim_{n\to\infty} v^n(x) = v(x)$ for all $x \in X$. If $\boldsymbol{a}^n \to \boldsymbol{a}^o \in A_t(x)$ as $n \to \infty$, then*

$$\lim_{n\to\infty} \sum_{y\in X} v^n(y)q_t(y|x,\boldsymbol{a}^n) = \sum_{y\in X} v(y)q_t(y|x,\boldsymbol{a}^o)$$

for all $x \in X$.

Proof. It is sufficient to show that $\boldsymbol{a} \to q_t(Y|x, \boldsymbol{a})$ is continuous on $A_t(x)$ for each $Y \subset X$ and $x \in X$. Let $Y = \{y_1, y_2, \ldots\}$ and $Y_n = \{y_1, \ldots, y_n\}$. Then $q_t(Y|x, \boldsymbol{a}) = \sup_{n\in\mathbb{N}} q_t(Y_n|x, \boldsymbol{a})$ is lower semicontinuous on $A_t(x)$. Similarly, $q_t(X \setminus Y|x, \boldsymbol{a})$ is also lower semicontinuous on $A_t(x)$. Since $q_t(Y|x, \boldsymbol{a}) = 1 - q_t(X \setminus Y|x, \boldsymbol{a})$, the function $q_t(Y|x, \cdot)$ is continuous on $A_t(x)$. Now the claim follows from Proposition 18 in Chapter 11 in Royden (1968). □

Lemma 3. *Under Assumption A for every $t \in \mathbb{N}$, $x_{t+1} \in X$ and $i_{t+1} \in G_{t+1}$ we have that*

$$\lim_{T\to\infty} \sup_{\boldsymbol{\varphi}_{[t+1,\infty)}\in\Phi_{[t+1,\infty)}} \|V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)}) - V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]})\| = 0.$$

Proof. Note that for any sequence of strategy profiles $\boldsymbol{\varphi}_{[t+1,\infty)} \in \Phi_{[t+1,\infty)}$, any $x_{t+1} \in X$ and $T \ge t+1$ we have

$$0 < V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]})(x_{t+1}) \le \bar{c} := \exp\left(\frac{-\alpha\beta rc}{1-\beta}\right), \tag{10}$$

where c is a constant in Assumption A. Choose any $\epsilon > 0$ and $T \ge t+1$ such that

$$r\alpha \sum_{\tau=T+1}^{\infty} \beta^{\tau-t} c > -\frac{\epsilon}{e\bar{c}}, \quad \text{where} \quad e = \exp(1).$$

Then, we have

$$\mathbb{E}_{x_{t+1}}^{\boldsymbol{\varphi}_{[t+1,\infty)}} \exp\left(r\alpha \sum_{\tau=t+1}^{T} \beta^{\tau-t} u_{i_\tau}(x_\tau, \boldsymbol{a}_\tau) + \frac{\epsilon}{e\bar{c}}\right)$$
$$\ge V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)})(x_{t+1})$$
$$\ge \mathbb{E}_{x_{t+1}}^{\boldsymbol{\varphi}_{[t+1,\infty)}} \exp\left(r\alpha \sum_{\tau=t+1}^{T} \beta^{\tau-t} u_{i_\tau}(x_\tau, \boldsymbol{a}_\tau) - \frac{\epsilon}{e\bar{c}}\right).$$

Hence,

$$V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]})(x_{t+1}) \exp\left(\frac{\epsilon}{e\bar{c}}\right) \ge V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)})(x_{t+1})$$
$$\ge V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]})(x_{t+1}) \exp\left(-\frac{\epsilon}{e\bar{c}}\right).$$

By (10) we finally get

$$-\epsilon < -\frac{\epsilon}{e} \le \bar{c}(e^{-\frac{\epsilon}{e\bar{c}}} - 1) \le$$
$$V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty]})(x_{t+1}) - V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]})(x_{t+1}) \le$$
$$\bar{c}(e^{\frac{\epsilon}{e\bar{c}}} - 1) < \epsilon.$$

Here, we used the fact that $e^s - 1 < es$ and $e^{-s} - 1 > -s$ for $s \in (0, 1)$. □

Lemma 4. *Under Assumption A for every $t \in \mathbb{N}$, $x_{t+1} \in X$ and $i_{t+1} \in G_{t+1}$ the function $V_{i_{t+1}}(\cdot)(x_{t+1})$ is continuous on $\Phi_{[t+1,\infty)}$.*

Proof. Assume that $\varphi_\tau^n \to \varphi_\tau^o$ in Φ_τ as $n \to \infty$ for every $\tau = t+1, \ldots, T$. In view of Lemma 3, it is enough to show that the function $V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]}^n)(x_{t+1})$ converges to $V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,T]}^o)(x_{t+1})$ as $n \to \infty$ for any $T \ge t+1$.

The proof proceeds by induction with respect to the length ℓ of a finite time horizon. For more convenient notation we set $x := x_{t+1}$ and $\boldsymbol{a} := \boldsymbol{a}_{t+1}$. If $\ell = 1$, then $T = t+1$ and claim follows from the convergence $\boldsymbol{\varphi}_{t+1}^n \to \boldsymbol{\varphi}_{t+1}^o$ in Φ_{t+1} and Assumption A, i.e.,

$$\int_{A_{t+1}(x)} \exp(r\alpha\beta u_{i_{t+1}}(x, \boldsymbol{a}))\boldsymbol{\varphi}_{t+1}^n(d\boldsymbol{a}|x) \to$$
$$\int_{A_{t+1}(x)} \exp(r\alpha\beta u_{i_{t+1}}(x, \boldsymbol{a}))\boldsymbol{\varphi}_{t+1}^o(d\boldsymbol{a}|x)$$

as $n \to \infty$ for all $x \in X$. Assume now that the claim holds for an ℓ-step game with $\ell = T - t - 1 \ge 1$, starting at $y = x_{t+2} \in X$. More precisely, we have

$$V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^n)(y) \to V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^o)(y) \quad \text{as } n \to \infty. \tag{11}$$

We show that the claim holds for $\ell + 1 = T - t$. Indeed, note that

$$\left| e^{r\alpha\beta u_{i_{t+1}}(x,\boldsymbol{a})} \left(\sum_{y\in X} V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^n)(y) q_{t+1}(y|x, \boldsymbol{a}) - \sum_{y\in X} V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^o)(y) q_{t+1}(y|x, \boldsymbol{a}) \right) \right| \le e^{-r\alpha\beta c} d^n(x),$$

where

$$d^n(x) := \sup_{\boldsymbol{a}\in A_{t+1}(x)} \sum_{y\in X} |V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^n)(y) - V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^o)(y)| q_{t+1}(y|x, \boldsymbol{a}).$$

By Lemma 2 and Assumption A for every $n \in \mathbb{N}$ there exists $\boldsymbol{a}^n \in A_{t+1}(x_{t+1})$ such that

$$d^n(x) = \sum_{y\in X} |V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^n)(y) - V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^o)(y)| q_{t+1}(y|x, \boldsymbol{a}^n).$$

Since $A_t(x)$ is compact, then there is a subsequence n' such that $\boldsymbol{a}^{n'}$ is convergent. Again Lemma 2 and (11) imply that $d^{n'}(x) \to 0$. Clearly, the latter fact is true for any subsequence n' of n for which $\boldsymbol{a}^{n'}$ is convergent. Hence,

$$\exp(r\alpha\beta u_{i_{t+1}}(x, \boldsymbol{a})) \sum_{y\in X} V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^n)(y) q_{t+1}(y|x, \boldsymbol{a}) \to$$
$$\exp(r\alpha\beta u_{i_{t+1}}(x, \boldsymbol{a})) \sum_{y\in X} V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^o)(y) q_{t+1}(y|x, \boldsymbol{a})$$

uniformly in $\boldsymbol{a}$ as $n \to \infty$. This fact and the convergence of $\boldsymbol{\varphi}_{t+1}^n \to \boldsymbol{\varphi}_{t+1}^o$ in Φ_{t+1} yield

$$\int_{A_{t+1}(x)} \exp(r\alpha\beta u_{i_{t+1}}(x, \boldsymbol{a})) \times \sum_{y\in X} V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^n)(y) q_{t+1}(y|x, \boldsymbol{a}) \boldsymbol{\varphi}_{t+1}^n(d\boldsymbol{a}|x) \to$$
$$\int_{A_{t+1}(x)} \exp(r\alpha\beta u_{i_{t+1}}(x, \boldsymbol{a})) \times \sum_{y\in X} V_{i_{t+2}}(\boldsymbol{\varphi}_{[t+2,T]}^o)(y) q_{t+1}(y|x, \boldsymbol{a}) \boldsymbol{\varphi}_{t+1}^o(d\boldsymbol{a}|x)$$

as $n \to \infty$ for every $x \in X$. This completes the proof. □

Lemma 5. *Let Assumption A hold. Assume that $\boldsymbol{\varphi}_{[t+1,\infty)}^n \to \boldsymbol{\varphi}_{[t+1,\infty)}^o$ in $\Phi_{[t,\infty)}$ as $n \to \infty$ for any $t \in \mathbb{N}$. Then for every $x_t \in X$ and $i_t \in G_t$*

$$\sup_{\boldsymbol{a}\in A_t(x_t)} P_{i_t}(x_t, \boldsymbol{\varphi}_{[t+1,\infty)}^n, \boldsymbol{a}) \to \sup_{\boldsymbol{a}\in A_t(x_t)} P_{i_t}(x_t, \boldsymbol{\varphi}_{[t+1,\infty)}^o, \boldsymbol{a}).$$

Proof. By Lemmas 2 and 4 we have that

$$V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)}^n)(x_{t+1}) \to V_{i_{t+1}}(\boldsymbol{\varphi}_{[t+1,\infty)}^o)(x_{t+1})$$

as $n \to \infty$ for every $x_{t+1} \in X$. Moreover, we have that

$$|\sup_{\boldsymbol{a}\in A_t(x_t)} P_{i_t}(x_t, \boldsymbol{\varphi}_{[t+1,\infty)}^n, \boldsymbol{a}) - \sup_{\boldsymbol{a}\in A_t(x_t)} P_{i_t}(x_t, \boldsymbol{\varphi}_{[t+1,\infty)}^o, \boldsymbol{a})|$$
$$\le \exp(-r\alpha c) D^n(x),$$

where

$D^n(x) :=$

$$\sup_{\boldsymbol{a}\in A_t(x_t)}\sum_{y\in X}\left|V_{i_{t+2}}(\boldsymbol{\varphi}^n_{[t+2,\infty)})(y)-V_{i_{t+2}}(\boldsymbol{\varphi}^o_{[t+2,\infty)})(y)\right| q_{t+1}(y|x,\boldsymbol{a}).$$

Now, analogously as in the proof of Lemma 4 we show that $D^n(x)\to 0$ as $n\to\infty$ for every $x\in X$ and the lemma follows. □

Let $\boldsymbol{\varphi}_{[1,\infty)} \in \Phi_{[1,\infty)}$ be any sequence of Markov strategy profiles for the players from G_t. By $\Phi_{[1,T]}$ we denote the set of sequences $(\boldsymbol{\varphi}_t)_{t\in[1,T]} = (\boldsymbol{\varphi}_t,\ldots,\boldsymbol{\varphi}_T)$.

Define

$$P_{i_t}(x_t,\boldsymbol{\varphi}_{[t+1,\infty)},\boldsymbol{\varphi}_t(x_t)) = \int_{A_t(x_t)} P_{i_t}(x_t,\boldsymbol{\varphi}_{[t+1,\infty)},\boldsymbol{a}_t)\boldsymbol{\varphi}_t(d\boldsymbol{a}_t|x_t),$$

for any $\boldsymbol{\varphi}_{[t,\infty)}\in\Phi_{[t,\infty)}$, where $P_{i_t}(x_t,\boldsymbol{\varphi}_{[t+1,\infty)},\boldsymbol{a}_t)$ is given in (8).

Definition 3. Let $\boldsymbol{\varphi}_{[1,\infty)}\in\Phi_{[1,\infty)}$ and $T\ge 1$ be fixed. A sequence $(\boldsymbol{\varphi}^*_t)_{t\in[1,T]}\in\Phi_{[1,T]}$ is a *Markov perfect equilibrium* in the T-step game with selves from the sets $G_1,\ldots,G_T$ ($MPE_{[1,T]}$), if for each $t=1,\ldots,T$, every self $i_t\in G_t$ and any Markov strategy $\xi_{i_t}\in\Phi_{i_t}$, it holds

$$P_{i_t}(x_t,\boldsymbol{\phi}_{[t+1,\infty)},\boldsymbol{\varphi}^*_t(x_t)) \ge P_{i_t}(x_t,\boldsymbol{\phi}_{[t+1,\infty)},[\boldsymbol{\varphi}^*_{-i_t},\xi_{i_t}](x_t))$$

for all $x_t\in X$, where $\boldsymbol{\phi}_k=\boldsymbol{\varphi}^*_k$ for $k=t+1,\ldots,T$ and $t<T$ and $\boldsymbol{\phi}_k=\boldsymbol{\varphi}_k$ for $k\ge T+1$.

It should be noted that in the game introduced here the players choose actions only on the first T steps. Next the fixed sequence $\boldsymbol{\varphi}_{[T+1,\infty)}$ is used. Hence, the game is called a T-step game. Moreover, Definition 3 says that for any $x_t\in X$, $t+1\le T$, $\boldsymbol{\varphi}^*_t(x_t)$ is a Nash equilibrium in the game $\Gamma(x_t,(\boldsymbol{\varphi}^*_{t+1},\ldots,\boldsymbol{\varphi}^*_T,\boldsymbol{\varphi}_{[T+1,\infty)}))$, and for any $x_T\in X$, $\boldsymbol{\varphi}^*_T(x_T)$ is a Nash equilibrium in the game $\Gamma(x_T,\boldsymbol{\varphi}_{[T+1,\infty)})$.

Proof of Theorem 1. Let us fix a sequence of Markov strategy profiles $\boldsymbol{\varphi}_{[1,\infty)}\in\Phi_{[1,\infty)}$. Now we show that there is an $MPE_{[1,T]}$ for any $T\ge 1$. The proof proceeds by backward induction.

Assume that $T=1$ and $x_1\in X$. Consider the one-shot game $\Gamma(x_1,\boldsymbol{\varphi}_{[2,\infty)})$ for the players from G_1. Recall that the payoff function for self $i_1\in G_1$ is $P_{i_1}(x_1,\boldsymbol{\varphi}_{[2,\infty)},\boldsymbol{a}_1)$, where $\boldsymbol{a}_1\in A_1(x_1)$ is a strategy profile used by the players from G_1. By Assumption A and Lemma 2 the function $P_{i_1}(x_1,\boldsymbol{\varphi}_{[2,\infty)},\cdot)$ is continuous on $A_1(x_1)$. By the Kakutani–Fan–Glicksberg fixed point theorem (Corollary 17.55 in Aliprantis and Border (2006)), there exists a Nash equilibrium $\boldsymbol{f}^{(1)}_1(x_1)$ in the game $\Gamma(x_1,\boldsymbol{\varphi}_{[2,\infty)})$ for every $x_1\in X$. Hence, $\boldsymbol{f}^{(1)}_1\in\Phi_1$ is a Nash equilibrium in the one-step model, i.e., $MPE_{[1,1]}$.

Now we consider the two-step model and construct an $MPE_{[1,2]}$. The players from G_2 in state $x_2\in X$ consider the one-shot game $\Gamma(x_2,\boldsymbol{\varphi}_{[3,\infty)})$ with the payoff function $P_{i_2}(x_2,\boldsymbol{\varphi}_{[3,\infty)},\boldsymbol{a}_2)$ for self $i_2\in G_2$ where $\boldsymbol{a}_2\in A_2(x_2)$. Again the Kakutani–Fan–Glicksberg fixed point theorem gives a Nash equilibrium $\boldsymbol{f}^{(2)}_2(x_2)$ for every $x_2\in X$ in the one-stage game $\Gamma(x_2,\boldsymbol{\varphi}_{[3,\infty)})$. Now, players from the set G_1 play the one-shot game $\Gamma(x_1,(\boldsymbol{f}^{(2)}_2,\boldsymbol{\varphi}_{[3,\infty)}))$ with the payoff function $P_{i_1}(x_1,(\boldsymbol{f}^{(2)}_2,\boldsymbol{\varphi}_{[3,\infty)}),\boldsymbol{a}_1)$ for self $i_1\in G_1$, $x_1\in X$. $\boldsymbol{a}_1\in A_1(x_1)$. This game has a Nash equilibrium $\boldsymbol{f}^{(2)}_1(x_1)$. In this way, we have obtained an $MPE_{[1,2]}$ of the form $(\boldsymbol{f}^{(2)}_1,\boldsymbol{f}^{(2)}_2)$ for the two-step game. Proceeding along the same lines, we can construct an $MPE_{[1,T]}$ of the form $(\boldsymbol{f}^{(T)}_1,\ldots,\boldsymbol{f}^{(T)}_T)$ for any number of steps $T\in\mathbb{N}$.

If we continue this procedure for longer and longer time horizons, we finally obtain an infinite matrix with the entries:

$$\begin{array}{ccccccccc}
\boldsymbol{f}^{(1)}_1 & \boldsymbol{\varphi}_2 & \boldsymbol{\varphi}_3 & \boldsymbol{\varphi}_4 & \cdots & \boldsymbol{\varphi}_t & \boldsymbol{\varphi}_{t+1} & \boldsymbol{\varphi}_{t+2} & \cdots\\
\boldsymbol{f}^{(2)}_1 & \boldsymbol{f}^{(2)}_2 & \boldsymbol{\varphi}_3 & \boldsymbol{\varphi}_4 & \cdots & \boldsymbol{\varphi}_t & \boldsymbol{\varphi}_{t+1} & \boldsymbol{\varphi}_{t+2} & \cdots\\
\boldsymbol{f}^{(3)}_1 & \boldsymbol{f}^{(3)}_2 & \boldsymbol{f}^{(3)}_3 & \boldsymbol{\varphi}_4 & \cdots & \boldsymbol{\varphi}_t & \boldsymbol{\varphi}_{t+1} & \boldsymbol{\varphi}_{t+2} & \cdots\\
\boldsymbol{f}^{(4)}_1 & \boldsymbol{f}^{(4)}_2 & \boldsymbol{f}^{(4)}_3 & \boldsymbol{f}^{(4)}_4 & \cdots & \boldsymbol{\varphi}_t & \boldsymbol{\varphi}_{t+1} & \boldsymbol{\varphi}_{t+2} & \cdots\\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots\\
\boldsymbol{f}^{(t)}_1 & \boldsymbol{f}^{(t)}_2 & \boldsymbol{f}^{(t)}_3 & \boldsymbol{f}^{(t)}_4 & \cdots & \boldsymbol{f}^{(t)}_t & \boldsymbol{\varphi}_{t+1} & \boldsymbol{\varphi}_{t+2} & \cdots\\
\boldsymbol{f}^{(t+1)}_1 & \boldsymbol{f}^{(t+1)}_2 & \boldsymbol{f}^{(t+1)}_3 & \boldsymbol{f}^{(t+1)}_4 & \cdots & \boldsymbol{f}^{(t+1)}_t & \boldsymbol{f}^{(t+1)}_{t+1} & \boldsymbol{\varphi}_{t+2} & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}$$

For each $t\in\mathbb{N}$, let $(\boldsymbol{\phi}^n_t)_{t\in\mathbb{N}}$ be the sequence in the nth row of the above matrix, i.e.,

$$\boldsymbol{\phi}^n_t := \boldsymbol{f}^{(n)}_t \text{ for } t\le n \quad\text{and}\quad \boldsymbol{\phi}^n_t=\boldsymbol{\varphi}_t \text{ for } t>n.$$

Since Φ_t is compact for every $t\in\mathbb{N}$, we may choose a subsequence $n_1(k)$ of the sequence (n) such that the entries $\boldsymbol{\phi}^{n_1(k)}_1$ in the first column of the matrix tend to some $\boldsymbol{\phi}^*_1\in\Phi_1$ as $n_1(k)\to\infty$. Now choose in the second column a subsequence $n_2(k)$ of $(n_1(k))$ such that $\boldsymbol{\phi}^{n_2(k)}_2$ converges to some $\boldsymbol{\phi}^*_2\in\Phi_2$ as $n_2(k)\to\infty$. Repeating this procedure, we infer that there exists a subsequence $(n_\tau(k))$ of $(n_{\tau-1}(k))$ such that $\boldsymbol{\phi}^{n_\tau(k)}_\tau$ converges to some $\boldsymbol{\phi}^*_\tau$ in Φ_τ. Consider the diagonal sequence $m(k):=n_k(k)$, $k\in\mathbb{N}$. Clearly, $\boldsymbol{\phi}^{m(k)}_t\to\boldsymbol{\phi}^*_t$ in Φ_t as $k\to\infty$ for every $t\in\mathbb{N}$. The limit sequence $(\boldsymbol{\phi}^*_t)_{t\in\mathbb{N}}$ represents an *MPE*. Indeed, choose any $t\in\mathbb{N}$ and observe that

$$P_{i_t}(x_t,\boldsymbol{\phi}^n_{[t+1,\infty)},\boldsymbol{\phi}^n_t(x_t)) \ge P_{i_t}(x_t,\boldsymbol{\phi}^n_{[t+1,\infty)},[\boldsymbol{\phi}^n_{-i_t},\xi_{i_t}](x_t))$$

for any $\xi_{i_t}\in\Phi_{i_t}$, $x_t\in X$ and $n\ge t$.

Take the subsequence $m(k)$ defined above and let $k\to\infty$. Then by Lemma 5

$$\begin{aligned}
&\lim_{k\to\infty} P_{i_t}(x_t,\boldsymbol{\phi}^{m(k)}_{[t+1,\infty)},\boldsymbol{\phi}^{m(k)}_t(x_t)) =\\
&P_{i_t}(x_t,\boldsymbol{\phi}^*_{[t+1,\infty)},\boldsymbol{\phi}^*_t(x_t)) \ge\\
&\lim_{k\to\infty} P_{i_t}(x_t,\boldsymbol{\phi}^{m(k)}_{[t+1,\infty)},[\boldsymbol{\phi}^{m(k)}_{-i_t},\xi_{i_t}](x_t)) =\\
&P_{i_t}(x_t,\boldsymbol{\phi}^*_{[t+1,\infty)},[\boldsymbol{\phi}^*_{-i_t},\xi_{i_t}](x_t))
\end{aligned}$$

for every $\xi_{i_t}\in\Phi_{i_t}$, $x_t\in X$. Now the claim easily follows from Lemma 1. □

Remark 4. Let us consider for a while a Markov decision process with a standard discounted payoff criterion, i.e., the case where every set G_t is a singleton and (3) is used with $\alpha=1$. The proof of Theorem 1 resembles the value iteration algorithm in discounted dynamic programming. Indeed, the rows in above matrix determine the iterative value functions converging to the value function in the infinite time horizon model, e.g., see Bertsekas and Shreve (1978) or Puterman (2005). It should be noted, however, that if $\alpha\ne 1$, then the above construction does not give a convergent sequence, see Example 5.6 in Jaśkiewicz and Nowak (2021). Hence, we believe in case of $\alpha\ne 1$, the same phenomenon occurs in the model involving entropic risk measures. Therefore, taking a subsequence as in the proof of Theorem 1 is necessary, even if each G_t is a singleton.

For a proof of Theorem 2 we recall that Φ_i is the set of all transition probabilities ϕ_i from X to A_i such that $\phi_i(A_i(x)|x)=1$ for all $x\in X$. Next we put $\Phi:=\prod_{i=1}^N\Phi_i$. An element of Φ is denoted by $\boldsymbol{\phi}=(\phi_1,\ldots,\phi_N)$. Let Φ_S be the set of all stationary sequences in $\Phi_{[1,\infty)}$. An element of Φ_S is of the form $\widehat{\boldsymbol{\phi}}=(\boldsymbol{\phi},\boldsymbol{\phi},\ldots)$. Note that $\widehat{\boldsymbol{\phi}}^n\to\widehat{\boldsymbol{\phi}}^o$ in Φ_S if and only if $\boldsymbol{\phi}^n\to\boldsymbol{\phi}^o$ in Φ.

If $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_N) \in \Phi$ and $\xi_i \in \Phi_i$, then $[\boldsymbol{\phi}_{-i}, \xi_i]$ is the strategy profile $\boldsymbol{\phi}$ with ϕ_i replaced by ξ_i. Then $[\boldsymbol{\phi}_{-i}, \xi_i](x) := (\psi_1(\cdot|x), \psi_2(\cdot|x), \ldots, \psi_N(\cdot|x))$ and

$$[\boldsymbol{\phi}_{-i}, \xi_i](d\boldsymbol{a}|x) := \psi_1(da_1|x) \otimes \psi_2(da_2|x) \otimes \cdots \otimes \psi_N(da_N|x)$$

with $\psi_j(\cdot|x) = \phi_j(\cdot|x)$ for $j \neq i$ and $\psi_i(\cdot|x) = \xi_i(\cdot|x)$. By definition, $[\boldsymbol{\phi}_{-i}, \xi_i](\cdot|x)$ is a product probability measure on $\prod_{i=1}^N A_i(x)$.

Proof of Theorem 2. For any $x \in X$ and $\boldsymbol{\phi} \in \Phi$ let

$$\widehat{V}_i(\widehat{\boldsymbol{\phi}})(x) := \mathbb{E}_x^{\widehat{\boldsymbol{\phi}}} \exp\left(\alpha \sum_{\tau=t+1}^{\infty} \beta^{\tau-t} r u_i(x_\tau, \boldsymbol{a}_\tau)\right). \tag{12}$$

Moreover, put

$$\begin{aligned}&\widehat{P}_i(x, \widehat{\boldsymbol{\phi}}, [\boldsymbol{\phi}_{-i}, \xi_i](x)) := \\ &-\int_{A(x)} \exp(r u_i(x, \boldsymbol{a})) \sum_{y \in X} V_i(\widehat{\boldsymbol{\phi}})(y) q(y|x, \boldsymbol{a}) [\boldsymbol{\phi}_{-i}, \xi_i](d\boldsymbol{a}|x).\end{aligned} \tag{13}$$

Let $\boldsymbol{\phi} \in \Phi$ and $\mathcal{R}_i(\boldsymbol{\phi})$ be the set of all $\tilde{\xi}_i \in \Phi_i$ such that

$$\widehat{P}_i(x, \widehat{\boldsymbol{\phi}}, [\boldsymbol{\phi}_{-i}, \tilde{\xi}_i](x)) \geq \widehat{P}_i(x, \widehat{\boldsymbol{\phi}}, [\boldsymbol{\phi}_{-i}, \xi_i](x))$$

for all $x \in X$, $\xi_i \in \Phi_i$.

Define $\mathcal{R}(\boldsymbol{\phi}) := \prod_{i=1}^N \mathcal{R}_i(\boldsymbol{\phi})$. Under our continuity Assumption A, every set $\mathcal{R}_i(\boldsymbol{\phi})$ is non-empty, compact and convex, thus so is $\mathcal{R}(\boldsymbol{\phi}) \subset \Phi$. Using our continuity Lemma 5, one can easily show that if $i \in \mathcal{N}$ and $\boldsymbol{\phi}^n \to \boldsymbol{\phi}^o$ in Φ as $n \to \infty$ and $\tilde{\xi}_i^n \in \mathcal{R}_i(\boldsymbol{\phi}^n)$ for all n and $\tilde{\xi}_i^n \to \tilde{\xi}_i^o$, then $\tilde{\xi}_i^o \in \mathcal{R}_i(\boldsymbol{\phi}^o)$. For proving this, the definition (13) is used, and it is important to note that $\boldsymbol{\phi}^n \to \boldsymbol{\phi}^o$ in Φ implies that $\widehat{\boldsymbol{\phi}}^n \to \widehat{\boldsymbol{\phi}}^o$ in Φ_S as $n \to \infty$. This property means that the correspondence $\boldsymbol{\phi} \to \mathcal{R}(\boldsymbol{\phi})$ is upper semicontinuous. By the Kakutani–Fan–Glicksberg fixed point theorem (Corollary 17.55 in Aliprantis and Border (2006)), there exists $\boldsymbol{\phi}^* \in \Phi$ such that $\boldsymbol{\phi}^* \in \mathcal{R}(\boldsymbol{\phi}^*)$. Consider any $t \in \mathbb{N}$ and note that $\widehat{\boldsymbol{\phi}}^* = \widehat{\boldsymbol{\phi}}^*_{[t+1,\infty)}$. Assume that $i = i_t$ and $x = x_t$. Then $\boldsymbol{\phi}^* \in \mathcal{R}(\boldsymbol{\phi}^*)$ implies that $\boldsymbol{\phi}^*(x_t)$ is a Nash equilibrium in the game $\Gamma(x_t, \widehat{\boldsymbol{\phi}}^*_{[t+1,\infty)})$. By Lemma 1, $\widehat{\boldsymbol{\phi}}^*$ is an *MPE* and since $\widehat{\boldsymbol{\phi}}^* \in \Phi_S$, it is an *SMPE*. □

5. Dynamic games on networks

In this section, we first consider a standard discounted stochastic game with $\alpha = 1$, $0 < \beta < 1$, and the time invariant utilities for all players. In this set-up, we can write $i_t = i$ for all $t \in \mathbb{N}$. Interesting applications of standard discounted stochastic games to operations research and engineering (e.g., queuing networks) can be found in Altman (1996) and Altman, Hordijk, and Spieksma (1997). Below we outline one application of our results to a class of dynamic games on networks.

Consider a dynamic game played by N players on a network. An undirected network is described by a symmetric $N \times N$ matrix $p = [p_{i,j}]$, with a convention that $p_{i,j} = p_{j,i} = 1$ if players i and j are connected and $p_{i,j} = p_{j,i} = 0$ if they are not. We conventionally assume that $p_{i,i} = 0$ for all $i \in \mathcal{N}$. Let P be a set of such matrices. Next, let $s \in S$ (S is a countable set) denote a state of nature. A state of nature determines the specific within-period game that is played. Then, a state of a game is summarized by $x = (p, s)$ and the set of such states is X. Note that X is countable.

At a state $x \in X$ and for a profile of actions $\boldsymbol{a} = (a_1, \ldots, a_N) \in A(x)$, the within-period payoff of player i is given by $u_i(x, \boldsymbol{a})$. Observe that u_i depends hence on both: a network p and the state of nature determining the game s to be played.

An action of player i, denoted by a_i, includes two components $a_i = (d_i, b_i)$, where $d_i \in D_i$ is player's i decision (e.g., effort) in the within-period game s on a given network p. We can assume that every set D_i is a compact interval in $\mathbb{R}_+$ and the profile of such decisions is denoted by $\boldsymbol{d} = (d_1, \ldots, d_N) \in D := \prod_{i=1}^N D_i$. A component $b_i = (b_i^j)_{j \neq i}$ is an $N-1$ dimensional vector and denotes player i's investment into establishing new connections on a network the next period, to maintain connections or to severe the connections in the network with the remaining players. Hence, b_i^j denotes an investment of player i to establish or maintain a connection with player j. For example, we can let $b_i^j \in I = [0, 1]$ for every $i, j \in \mathcal{N}$ and $j \neq i$. The profile of such decisions $(b_1, \ldots, b_N)$ is denoted by $\boldsymbol{b}$.

The network evolution process is random, however, in each period some connections can be broken and new connections can be established. The probability of a network p' to be drawn (the next period) from P depends on selected profiles $\boldsymbol{b}$ and $\boldsymbol{d}$ and a state x. The joint probability distribution on $X = P \times S$ is summarized by $q(\cdot|x, \boldsymbol{a})$.

Some examples of stochastic network formation processes can be found, e.g., in Jackson and Watts (2002), who discuss problems arising when studying farsighted players in dynamic network formation games. These examples inspired our approach, but the class of games we study is slightly different. Models of interactions on a network embrace a wide variety of applications to the problems of, e.g., financing of local public goods, team organization, buyer–seller transactions on networks or labor market contracts, see Jackson (2005) and Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2010) and the references cited therein.

Below we give an illustration of the payoffs u_i and transition q that can be used in such applied models.

Example 1. The within-period payoff of player i for a profile $\boldsymbol{a} = (\boldsymbol{d}, \boldsymbol{b})$ at state $x = (p, s)$ is given, for example, by: $u_i(x, \boldsymbol{a}) \equiv u_i(p, s, \boldsymbol{d}, \boldsymbol{b}) := \pi_i(x, \boldsymbol{d}) - \gamma_i(x, b_i)$, where $\pi_i : X \times D \to \mathbb{R}$ with $\pi_i(x, \boldsymbol{d})$ denoting the instantaneous payoff of playing game s on p, and $\gamma_i : X \times I^{N-1} \to \mathbb{R}_+$ with $\gamma_i(x, b_i)$ denoting the cost of investment b_i in the network formation process.

In a game of local interactions with neighbors on $p \in P$, the set of neighbors of player i is given by $N_i(p) = \{j \in \{1, \ldots, N\} : p_{i,j} = 1\}$ and its cardinality is denoted by $\overline{N}_i(p)$. An example of payoff π_i can be taken as:

$$\pi_i(x, \boldsymbol{d}) = \pi_i(p, s, \boldsymbol{d}) = h\left(d_i + s \frac{\sum_{j \in N_i(p)} d_j}{\overline{N}_i(p)}\right) - c_i(d_i),$$

where $h : \mathbb{R} \to \mathbb{R}_+$ is a non-decreasing, convex function and $c_i : \mathbb{R}_+ \to \mathbb{R}_+$ is a cost function associated with own action d_i (e.g. effort). Here, the state s determines the nature of the externalities. For instance, $S = \{-1, 1\}$ and $s = 1$ means that players play a coordination game with local positive externalities (from their neighbors), but $s = -1$ means that the players play a game with negative local externalities (from their neighbors). Finally, $\gamma_i(x, b_i)$ denotes the cost of investing into a network formation at state x. For example, a cost of investment can be lower, if many connections already exist.

An example of a transition probability between state x and $x' = (p', s')$ is $q(x'|x, \boldsymbol{a}) = \lambda(p'|p, \boldsymbol{b}) \otimes \nu(s')$, where $\nu(s') = 0.5$ for $s' = -1$ and $s' = 1$. Clearly, the state of nature s' can evolve stochastically in a more complicated version, than the specified above. Namely, its transition can depend on the current network p as well as current actions.

The transition probability λ from p to a new network $p' = [p'_{i,j}]$, $i, j \in \mathcal{N}$, can be constructed as follows: a link between players i and $j \neq i$ will be added or maintained (i.e., $p'_{i,j} = 1$) with probability $0.5(1 - p_{i,j}) b_i^j b_j^i + p_{i,j} b_i^j b_j^i$. That is, if the link between i and j does not exist ($p_{i,j} = 0$) it can be created, if players jointly invest in its creation (i.e. $b_i^j > 0$ and $b_j^i > 0$) and if it does exist ($p_{i,j} = 1$) it may be maintained by such a joint investment. Observe that for a given vector $\boldsymbol{b}$ the probability of creating a new link between i and j is lower than the probability of successfully maintaining an already existing link. Moreover, if any of the players (for example player i) does not want to establish a new link with player j, or wants to delete the existing link, it can choose $b_i^j = 0$ and hence make a probability of $p'_{i,j} = 1$ equal to zero. Clearly, we have $p'_{i,j} = p'_{j,i}$.

Since the network evolution process and the states of nature are random, the players form expectations about the future trajectories of states and actions. In our setting, they use the certainty equivalent of the exponential function to calculate their payoffs in the infinite time horizon case with a long-run discount factor β and short-run discount factor α. Their payoffs are given in (2). If $\alpha = 1$ a Nash equilibrium in Markov strategies then exists by the work of Basu and Ghosh (2018). However, this equilibrium is not time-consistent. The time-consistent solution follows from Theorem 2 that says that this game possesses an *SMPE*.

References

Aliprantis, C. D., & Border, K. C. (2006). *Infinite dimensional analysis. A hitchhiker's guide*. New York: Springer.

Alj, A., & Haurie, A. (1983). Dynamic equilibria in multigeneration stochastic games. *IEEE Transactions on Automatic Control*, *28*, 193–203.

Altman, E. (1996). Nonzero-sum stochastic games in admission, service and routing control in queueing systems. *Queueing Systems*, *23*, 259–279.

Altman, E., Hordijk, A., & Spieksma, F. M. (1997). Contraction conditions for average and α-discount optimality in countable state Markov games with unbounded rewards. *Mathematics of Operations Research*, *22*, 588–618.

Balbus, Ł., Jaśkiewicz, A., & Nowak, A. S. (2020). Markov perfect equilibria in a dynamic decision model with quasi-hyperbolic discounting. *Annals of Operations Research*, *287*, 573–591.

Balbus, Ł., & Nowak, A. S. (2008). Existence of perfect equilibria in a class of multigenerational stochastic games of capital accumulation. *Automatica*, *44*, 1471–1479.

Balbus, Ł., Reffett, K., & Woźny, Ł. (2023). Time consistent policies and quasi-hyperbolic discounting. In A. Banerjee (Ed.), *Oxford research encyclopedia of economics and finance*. Oxford, UK: Oxford University Press.

Başar, T. (1999). Nash equilibria of risk-sensitive nonlinear stochastic differential games. *Journal of Optimization Theory and Applications*, *100*, 479–498.

Başar, T. (2021). Robust designs through risk sensitivity: an overview. *Journal of Systems Science and Complexity*, *34*, 1634–1665.

Basu, A., & Ghosh, M. K. (2018). Nonzero-sum risk-sensitive stochastic games on a countable state space. *Mathematics of Operations Research*, *43*, 516–532.

Bäuerle, N., & Jaśkiewicz, A. (2015). Risk-sensitive dividend problems. *European Journal of Operational Research*, *242*, 161–171.

Bäuerle, N., & Jaśkiewicz, A. (2024). Markov decision processes with risk-sensitive criteria: an overview. *Mathematical Methods of Operations Research*, *99*, 141–178.

Bäuerle, N., & Rieder, U. (2014). More risk-sensitive Markov decision processes. *Mathematics of Operations Research*, *39*, 105–120.

Bertsekas, D. P., & Shreve, S. E. (1978). *Stochastic optimal control the discrete time case*. New York: Academic Press.

Björk, T., Khapko, M., & Murgoci, A. (2021). *Time-inconsistent control theory with finance applications*. Springer Cham.

Caravani, P., & Papavassilopoulos, G. (1990). A class of risk-sensitive noncooperative games. *Journal of Economic Dynamics & Control*, *14*, 117–149.

Christensen, S., & Lindensjö, K. (2018). On finding equilibrium stopping times for time-inconsistent Markovian problems. *SIAM Journal on Control and Optimization*, *56*, 4228–4255.

Chung, K.-J., & Sobel, M. J. (1987). Discounted MDP's: Distribution functions and exponential utility maximization. *SIAM Journal on Control and Optimization*, *25*, 49–62.

Di Masi, G. B., & Stettner, Ł. (2000). Infinite horizon risk sensitive control of discrete time Markov processes with small risk. *Systems & Control Letters*, *40*, 15–20.

Escobar, J. F. (2013). Equilibrium analysis of dynamic models of imperfect competition. *International Journal of Industrial Organization*, *31*(1), 92–101.

Eshwar, S. R., Motwani, M., Roy, N., & Thoppe, G. (2024). Reinforcement learning with quasi-hyperbolic discounting (extended abstract). In *2024 tenth Indian control conference* (pp. 512–513). Bhopal, India: IEEE.

Föllmer, H., & Schied, A. (2004). *Stochastic finance*. Berlin: Walter de Gruyter.

Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., & Yariv, L. (2010). Network games. *Review of Economic Studies*, *77*(1), 218–244.

Hansen, L. P., & Sargent, T. J. (1995). Discounted linear exponential quadratic Gaussian control. *IEEE Transactions on Automatic Control*, *40*, 968–971.

Harris, C., & Laibson, D. (2001). Dynamic choices of hyperbolic consumers. *Econometrica*, *69*, 935–957.

Haurie, A. (2005). A multigenerational game model to analyze sustainable development. *Annals of Operations Research*, *137*, 369–386.

Haurie, A., & Moresino, F. (2006). Computing equilibria in stochastic games of intergenerational equity. *International Game Theory Review*, *8*, 273–293.

Howard, R. A., & Matheson, J. E. (1972). Risk-sensitive Markov decision processes. *Management Science*, *18*, 356–369.

Jackson, M. O. (2005). A survey of network formation models: stability and efficiency. In G. Demange, & M. Wooders (Eds.), *Group formation in economics: networks, clubs, and coalitions* (pp. 11–57). Cambridge University Press.

Jackson, M. O., & Watts, A. (2002). The evolution of social and economic networks. *Journal of Economic Theory*, *106*, 265–295.

Jacobson, D. (1973). Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control*, *18*, 124–131.

Jaquette, S. C. (1973). Markov decision processes with a new optimality criterion: Discrete time. *The Annals of Statistics*, *1*, 496–505.

Jaquette, S. C. (1976). A utility criterion for Markov decision processes. *Management Science*, *23*, 43–49.

Jaśkiewicz, A., & Nowak, A. S. (2014). Stationary Markov perfect equilibria in risk sensitive stochastic overlapping generations models. *Journal of Economic Theory*, *151*, 411–447.

Jaśkiewicz, A., & Nowak, A. S. (2021). Markov decision processes with quasi-hyperbolic discounting. *Finance and Stochastics*, *25*, 189–229.

Jaśkiewicz, A., & Nowak, A. S. (2024). Markov perfect equilibria for risk-sensitive Markov decision processes with quasi-hyperbolic discounting. *Discrete and Continuous Dynamical Systems-B*.

Klompstra, M. B. (2000). Nash equilibria in risk-sensitive dynamic games. *IEEE Transactions on Automatic Control*, *45*, 1397–1401.

Krusell, P., Kuruş¸cu, B., & Smith, A. (2002). Equilibrium welfare and government policy with quasi-geometric discounting. *Journal of Economic Theory*, *105*, 42–72.

Krusell, P., & Smith, A. (2003). Consumption–savings decisions with quasi–geometric discounting. *Econometrica*, *71*, 365–375.

Phelps, E., & Pollak, R. (1968). On second best national savings and game equilibrium growth. *Review of Economic Studies*, *35*, 195–199.

Pitera, M., & Stettner, Ł. (2023). Discrete-time risk sensitive portfolio optimization with proportional transaction costs. *Mathematical Finance*, *33*, 1287–1313.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, *32*, 122–136.

Puterman, M. L. (2005). *Markov decision processes: discrete stochastic dynamic programming*. New York: Wiley and Sons.

Royden, H. (1968). *Real analysis*. New York: Macmillan.

Sargent, T. J., & Stachurski, J. (2023). *Dynamic programming. Volume I: finite states*. Foundations Quantecon.

Shapley, L. (1953). Stochastic games. *Proceedings of the National Academy of Sciences of the USA*, *39*, 1095–1100.

Solano, J. M. (2014). Time-consistent equilibria in a differential game model with time inconsistent preferences and partial cooperation. In J. Haunschmied, V. Veliov, & S. Wrzaczek (Eds.), *Dynamic games in economics. dynamic modeling and econometrics in economics and finance*: *Vol. 16*, (pp. 219–238). Berlin, Heidelberg: Springer.

Solano, J. M., & Shevkoplyas, E. V. (2011). Non-constant discounting and differential games with random time horizon. *Automatica*, *47*, 2626–2638.

Strotz, R. H. (1956). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies*, *23*(3), 165–180.

Whittle, P. (1990). *Risk-sensitive optimal control*. New York: Wiley.

Anna Jaśkiewicz received a Ph.D. degree in Mathematics from Wrocław University of Technology (Poland) in 2003. Her research interests include game theory, stochastic dynamic programming, and mathematical economics. Currently she is a professor at the Faculty of Pure and Applied Mathematics at Wrocław University of Science and Technology in Poland. Currently she serves as an Associate Editor for Dynamic Games and Applications, SIAM Journal on Control and Optimization and Operations Research Letters.

Andrzej S. Nowak received Ph.D. degree in Mathematics from Wrocław University of Technology (Poland) in 1981. His research activities include game theory, stochastic dynamic optimization, and mathematical economics. Currently he is a professor at the Mathematics Department of the University of Zielona Góra in Poland. In the past he was an Associate Editor for International Journal of Game Theory, Mathematical Methods of Operations Research, Dynamic Games and Applications. At present he belongs to the Advisory Board of Dynamic Games and Applications. During the 18th International Symposium on Dynamic Games and Applications in Grenoble (France) in July, 2018, Andrzej S. Nowak received the Rufus Isaacs Award for his contribution to stochastic dynamic game theory.

Łukasz Woźny received a Ph.D. degree in Economics from SGH Warsaw School of Economics (Poland) in 2008. His research interests include game theory and mathematical economics, especially time consistency, general equilibrium, and recursive competitive equilibrium. He is a professor at the Collegium of Economic Analysis at the SGH Warsaw School of Economics in Poland. Currently he serves as an Co-Editor of Statistical Review.