APPLICATIONES MATHEMATICAE
50,2 (2023), pp. 107–121

Łukasz Balbus (Zielona Góra)
Anna Jaśkiewicz (Wrocław)
Andrzej S. Nowak (Zielona Góra)

NON-ZERO-SUM STOCHASTIC GAMES WITH RECURSIVE UTILITIES OF RISK-SENSITIVE PLAYERS

Abstract. Recursive utilities constructed by conditional entropic risk measures have recently been considered in various stochastic models and their applications, e.g., in economic dynamics. We study countable state discounted stochastic games played by risk-sensitive players. More precisely, we assume that the players evaluate their payoffs in a recursive way with the aid of the certainty equivalent of an exponential utility function. Under typical continuity and compactness conditions we prove that a stationary Nash equilibrium exists.

1. Introduction. The seminal papers of Howard and Matheson [22], Jacobson [24] and Jaquette [25, 26] mark the beginning of highly active research on dynamic programming with risk-sensitive preferences of the controller. While maximising the expected payoff implies a risk-neutral attitude, empirical evidence suggests that many agents tend to be risk-averse or are even made to be so by regulations, e.g. in finance or insurance industry [30]. The term "a risk-sensitive decision maker" mainly refers in the literature to the situation when the expectation of the random payoff is replaced by the certainty equivalent of an exponential utility. Thus, such a risk-sensitive model assumes that a controller uses a non-linear utility function. It is worth mentioning that the negative of the certainty equivalent of an exponential utility is also known as the entropic risk measure [20].

2020 *Mathematics Subject Classification*: Primary 91A15; Secondary 91A10, 90C39, 90C40, 91G70.
Key words and phrases: recursive utility, dynamic programming, risk-sensitive player, stationary Nash equilibrium.
Received 19 September 2023; revised 18 December 2023.
Published online 7 February 2024.

In recent years, there has been intensive development in this direction. For example, the approach based on the exponential utility found useful applications, in economics [2, 10, 27, 31, 36, 39], in actuarial science [9], in finance [20, 33] and in operations research [21, 24, 40, 41]. Moreover, as argued by Hansen and Sargent [21] and Başar [6] the risk-sensitive preferences are also attractive, because they can be used to model preferences for robustness. In such a case, one can interpret the risk parameter in the exponential utility function as the robustness parameter. Furthermore, the risk-sensitive preferences are also interesting from a mathematical point of view and have inspired a stream of works in Markov decision processes (for example, [3, 11, 16, 17]) and in dynamic games, see [7, 12] for zero-sum games and [8, 5, 15, 28, 32, 38] for non-zero-sum games.

In this paper, we consider a non-zero-sum stochastic game on a countable state space with the risk-averse players. In other words, the players are equipped with a parameter that reflects their risk-averse attitude towards risk. In the literature, there are two approaches that use the certainty equivalent of an exponential function. The first one corresponds to the case when it is defined on the space of all infinite histories (plays, say Ω) in the game or decision processes with the measure $\mathbb{P}$ constructed by strategies of the agents and transition probability according to the Ionescu–Tulcea theorem (see details in Subsection 3.2). The random variable is then the discounted payoff defined on Ω. The second method, on the other hand, is related to the use of the certainty equivalent of an exponential function sequentially, i.e., at every step. The controller accepts in each period a certainty equivalent as a terminal payment instead of continuing the dynamic choice process. This technique leads to the so-called *recursive utilities*. These two frameworks have been developed in parallel. The reader is referred to [11, 16, 17] and to [3, 9, 10, 21] where the former and latter approaches, respectively, were examined for Markov decision models with discounted payoffs.

The objective of our paper is to apply the latter method for stochastic games on a countable state space. The certainty equivalent of an exponential function is used sequentially to define discounted *recursive utilities* for the players. More precisely, in the nth step of the game, it is a probability measure on the product of the state space and the set of action profiles of the players. It depends on the history up to the nth state. This method leads to stationary Nash equilibria in discounted stochastic games under consideration. The fact is in contrast to [8], where the authors obtained a Nash equilibrium in the class of Markovian strategies for the discounted payoffs defined as in the first method. To the best of our knowledge, there is only one paper [2] on a dynamic game, where discounted *recursive utilities* are studied. This work is devoted to a special model of a two-player symmetric resource extraction game with specific transition probabilities. In [2] the authors showed that there exists a symmetric Nash equilibrium in the class of stationary strategies. Finally, recursive utilities were also employed to the intergenerational games [4, 27], in which the generations (players or selves) are risk-averse. However, these models are of different nature than the typical stochastic game. Finally, we wish to emphasise that dynamic games with related payoff criteria such as a weighted sum of mean and variance of the random payoffs need not possess Nash equilibria. Section 3.3 includes a simple example of a dynamic game with players who evaluate their payoffs as in the Markowitz model. We prove that the game has no Nash equilibrium.

2. Preliminaries and notation. We use $\mathbb{R}$ and $\mathbb{N}$ to denote the sets of all real numbers and positive integers, respectively. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let Z be a random *payoff* of a player on Ω that is essentially bounded. Instead of using $\mathbb{E}Z$ the player applies the exponential utility function to evaluate his/her random payoff. In what follows let $\mathcal{U}(z) = re^{rz}$ where $r \in \mathbb{R} \setminus \{0\}$. The *certainty equivalent* of $\mathcal{U}$ for the random variable Z is the number $c_e(r, Z)$ such that $\mathcal{U}(c_e(r, Z)) = \mathbb{E}\mathcal{U}(Z)$. It is easy to see that

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

The quantity $-c_e(r, Z)$ is also known as the entropic risk measure of Z (see [20]). However, in what follows we shall refer to $c_e(r, Z)$ as the entropic risk measure. The parameter $-r$ is known as the Arrow–Pratt risk coefficient of absolute risk aversion of $\mathcal{U}$ (see [20, p. 74] and [34]). Here, we call it simply the *risk sensitivity coefficient* of a player. A player equipped with the risk sensitivity coefficient $-r > 0$ (resp. $-r < 0$) is risk-averse (resp. risk-seeking) and indifferent between receiving a random payoff Z and obtaining the amount $c_e(r, Z)$ for sure. Observe that by applying the Taylor expansion around $r = 0$ for $\mathcal{U}$ we get

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

Therefore, if $r < 0$, then the individual who considers $c_e(r, Z)$ tends not only the maximisation of the expected value $\mathbb{E}Z$ of the random payoff Z, but also to minimisation of its variance.

Let K be a compact metric space. We denote by $C(K)$ the Banach space of all continuous real-valued functions on K endowed with the maximum norm. By the Riesz representation theorem [1, Corollary 14.15] the topological dual space $C^*(K)$ of $C(K)$ consists of all finite countably additive regular measures on the Borel sets in K. The space $C^*(K)$ is endowed with the weak-star metrisable topology; it is a locally convex linear topological space. The set $\Pr(K)$ of all probability measures on K is a compact convex subset of $C^*(K)$ [1, Theorem 6.21]. Recall that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of probability measures in $\Pr(K)$ weak-star converges to μ (denoted by $\mu_n \to^* \mu$) if $\lim_{n\to\infty} \int_K f\, d\mu_n = \int_K f\, d\mu$ for each $f \in C(K)$ (see [13, Proposition 7.21]).

If K is a metric space, then we denote by $B(K)$ 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 result. Fix $N \geq 2$. An N-person non-zero-sum *discounted stochastic game* (DSG) is defined by the following objects:

- $\mathcal{N} = \{1, \ldots, N\}$ is the *set of players.*
- X is a countable *state space*, endowed with the σ-algebra of all its subsets.
- A_i is a compact metric *action space* for player $i \in \mathcal{N}$, endowed with the Borel σ-algebra. For $x \in X$ the set $A_i(x)$ is a non-empty *compact* subset of A_i and denotes the set of admissible actions for player $i \in \mathcal{N}$ in x. We define
$$A := \prod_{i=1}^{N} A_i \quad \text{and} \quad A(x) := \prod_{i=1}^{N} A_i(x).$$
Note that
$$\mathbb{K}_i = \{(x, a_i) : x \in X,\ a_i \in A_i(x)\}$$
is the set all of *feasible state-action pairs* for player $i \in \mathcal{N}$. We also define
$$\mathbb{K} := \{(x, \boldsymbol{a}) : x \in X,\ \boldsymbol{a} = (a_1, \ldots, a_n) \in A(x)\}.$$
- $u_i : \mathbb{K} \to \mathbb{R}$, $i \in \mathcal{N}$, is a *utility-per-stage function* for player $i \in \mathcal{N}$.
- $q(y|x, \boldsymbol{a})$ is the transition probability from x to $y \in X$, when the players choose a profile $\boldsymbol{a} = (a_1, \ldots, a_N)$ of actions in $A(x)$.
- $\beta \in (0, 1)$ is the *discount factor.*
- $-r > 0$ is the *risk sensitivity coefficient* of each player $i \in \mathcal{N}$.

We impose the following assumptions.

Assumption **A**.

(i) *The functions $u_i(x, \cdot)$ are non-negative, bounded and continuous on $A(x)$ for all $x \in X$ and $i \in \mathcal{N}$.*

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

Let $H^1 = X$ and $H^{t+1} = \mathbb{K} \times H^t$ for $t \in \mathbb{N}$. Assume that every set H^t is endowed with its Borel σ-algebra. Clearly, $h^1 = x^1$ and an element $h^t = (x^1, \boldsymbol{a}^1, \ldots, x^t)$ of H^t represents the history of the game up to the tth stage, where $\boldsymbol{a}^k = (a_1^k, \ldots, a_N^k)$ is the profile of actions chosen by the players in state x^k at stage $k \in \mathbb{N}$ of the game.

A *strategy* for player $i \in \mathcal{N}$ is a sequence $\sigma_i = (\sigma_i^t)_{t\in\mathbb{N}}$, where each σ_i^t is a transition probability from H^t to A_i such that $\sigma_i^t(A_i(x^t)|h^t) = 1$ for any history $h^t \in H^t$, $t \in \mathbb{N}$. We denote by Σ_i the *set of all strategies* for player i. We let Φ_i denote the set of transition probabilities from X to A_i.

Then $\varphi_i \in \Phi_i$ if $\varphi_i(A_i(x)|x) = 1$ for all $x \in X$. A *stationary strategy* for player i is a constant sequence $(\varphi_i^t)_{t\in\mathbb{N}}$, where $\varphi_i^t = \varphi_i$ for all $t \in \mathbb{N}$ and some $\varphi_i \in \Phi_i$. For convenience, we shall identify a stationary strategy $(\varphi_i, \varphi_i, \ldots)$ for player i with the constant element φ_i of the sequence. Thus, the *set of all stationary* strategies of player i will also be denoted by Φ_i. Furthermore, we set

$$\Sigma = \prod_{i=1}^{N} \Sigma_i \quad \text{and} \quad \Phi = \prod_{i=1}^{N} \Phi_i.$$

Hence, Σ (resp. Φ) is the set of all (resp. all stationary) strategy profiles of the players.

3.1. DSGs with recursive utilities involving entropic risk measures. Assume that $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma$ where $\sigma_i = (\sigma_i^t)_{t\in\mathbb{N}}$ and $\boldsymbol{\sigma}^t := (\sigma_1^t, \ldots, \sigma_N^t)$. Then $\boldsymbol{\sigma}^t(d\boldsymbol{a}|h^t) := \sigma_1^t(da_1|h^t) \otimes \cdots \otimes \sigma_N^t(da_N|h^t)$ is the product probability measure on $A(x^t)$ induced by $\sigma_j^t(da_j|h^t)$, $j \in \mathcal{N}$. Let $v_i \in B(H^{t+1})$ and $h^t \in H^t$. We define

$$\text{(3.1)} \quad T^i_{\boldsymbol{\sigma}^t} v_i(h^t) := \frac{1}{r} \ln \int_{A(x^t)} \sum_{x^{t+1}\in X} e^{r(u_i(x^t,\boldsymbol{a})+\beta v_i(h^t,\boldsymbol{a},x^{t+1}))} q(x^{t+1}|x^t,\boldsymbol{a})\, \boldsymbol{\sigma}^t(d\boldsymbol{a}|h^t).$$

Note that, if $v_i, w_i \in B(H^{t+1})$ and $v_i \le w_i$ and $c \in \mathbb{R}$, then for $h^t \in H^t$,

$$\text{(3.2)} \qquad T^i_{\boldsymbol{\sigma}^t} v_i(h^t) \le T^i_{\boldsymbol{\sigma}^t} w_i(h^t) \quad \text{and} \quad T^i_{\boldsymbol{\sigma}^t}(v_i + c)(h^t) = T^i_{\boldsymbol{\sigma}^t} v_i(h^t) + \beta c.$$

REMARK 3.1. The operator in (3.1) is defined using the entropic risk measure induced by the probability measure $\mathbb{P}(x^{t+1}, d\boldsymbol{a}) = q(x^{t+1}|x^t,\boldsymbol{a})\boldsymbol{\sigma}^t(d\boldsymbol{a}|h^t)$, $h^t \in H^t$ on $\Omega = X \times A$, depending in step t on h^t. This risk measure acts as an aggregator of the t-stage utility $u_i(x^t,\cdot)$ and the discounted value $v_i(\cdot)$ of utilities to be received from stage $t+1$ onwards. In the risk-neutral case ($r = 0$) the operator in (3.1) takes on the following well-known form:

$$\widehat{T}^i_{\boldsymbol{\sigma}^t} v_i(h^t) = \int_{A(x^t)} \Big(u_i(x^t,\boldsymbol{a}) + \beta \sum_{x^{t+1}\in X} v_i(h^t,\boldsymbol{a},x^{t+1}) q(x^{t+1}|x^t,\boldsymbol{a}) \Big) \boldsymbol{\sigma}^t(d\boldsymbol{a}|h^t)$$

for any $h^t \in H^t$. A non-zero-sum *stochastic* game with a standard discounted payoff criterion and a countable state space was studied in [19, 37] and [18, 23]. In particular, Fink [19] and Takahashi [37] proved independently the existence of a stationary Nash equilibrium for the finite state case. Their result was extended to possibly infinite countable state space models by Federgruen [18].

Assume that $\mathbf{0} \equiv 0$ is the *null function*. Clearly $\mathbf{0} \in B(H^{t+1})$. By an *n-stage game* we mean a game with the *n-stage utility* for player i under a strategy profile $\boldsymbol{\sigma} \in \Sigma$ defined as

$$U_i^{[1,n]}(\boldsymbol{\sigma})(x^1) := T^i_{\boldsymbol{\sigma}^1} \cdots T^i_{\boldsymbol{\sigma}^n}\mathbf{0}(x^1), \quad x^1 \in X.$$

From (3.2) and Assumption **A** it follows that

$$U_i^{[1,n+1]}(\boldsymbol{\sigma})(x^1) = T^i_{\boldsymbol{\sigma}^1} \cdots T^i_{\boldsymbol{\sigma}^{n+1}}\mathbf{0}(x^1) \geq T^i_{\boldsymbol{\sigma}^1} \cdots T^i_{\boldsymbol{\sigma}^n}\mathbf{0}(x^1) = U_i^{[1,n]}(\boldsymbol{\sigma})(x^1).$$

Let $L > 0$ be such that $0 \leq u_i \leq L$ for all $i \in \mathcal{N}$. Using Jensen's inequality n times we obtain

$$0 \leq U_i^{[1,n]}(\boldsymbol{\sigma})(x^1) \leq L(1 + \beta + \cdots + \beta^{n-1}).$$

Therefore, our next definition is correct. The *recursive utility* for player i under a strategy profile $\boldsymbol{\sigma} \in \Sigma$ is

$$U_i(\boldsymbol{\sigma})(x^1) := \lim_{n\to\infty} U_i^{[1,n]}(\boldsymbol{\sigma})(x^1), \quad x^1 \in X. \tag{3.3}$$

As usual, for any $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_N) \in \Sigma$, $i \in \mathcal{N}$ and $\pi_i \in \Sigma_i$, we denote by $[\boldsymbol{\sigma}_{-i}, \pi_i]$ the strategy profile where player i uses π_i and every player $j \in \mathcal{N} \setminus \{i\}$ uses σ_j. We identify $[\boldsymbol{\sigma}_{-i}, \sigma_i]$ with $\boldsymbol{\sigma}$.

Definition 3.2. A strategy profile $\boldsymbol{\sigma}^* \in \Sigma$ is a *Nash equilibrium* in the DSG if

$$U_i(\boldsymbol{\sigma}^*)(x^1) \geq U_i([\boldsymbol{\sigma}^*_{-i}, \sigma_i])(x^1)$$

for every $\sigma_i \in \Sigma_i$, every player $i \in \mathcal{N}$ and for all $x^1 \in X$.

Theorem 3.3 (The Equilibrium Theorem). *Under Assumption* **A**, *the DSG with recursive utilities of risk-sensitive players has a stationary Nash equilibrium* $\boldsymbol{\varphi}^* \in \Phi$.

Remark 3.4. Our assumption that the functions u_i, $i \in \mathcal{N}$, are non-negative is not restrictive. If u_i is bounded and not necessarily $u_i \geq 0$, then we can find a constant $c > 0$ such that $u_i^+ = u_i + c \geq 0$ for all i. If $U_i^+(\boldsymbol{\sigma})(x^1)$ is the recursive utility of player i in the game with one-stage utility function u_i^+, then we have

$$U_i^+(\boldsymbol{\sigma})(x^1) = U_i(\boldsymbol{\sigma})(x^1) + \frac{c}{1-\beta}.$$

Therefore, we may restrict attention to games with non-negative rewards.

Remark 3.5. If we assume that $\boldsymbol{\sigma} = \boldsymbol{\varphi} \in \Phi$, then $\boldsymbol{\sigma}^t = \boldsymbol{\varphi}$ for all t and for $v_i \in B(X)$ formula (3.1) has a simpler form:

$$\begin{aligned} T^i_{\boldsymbol{\sigma}^t} v_i(h^t) &= T^i_{\boldsymbol{\varphi}} v_i(x^t) \\ &= \frac{1}{r} \ln \int_{A(x^t)} \sum_{x^{t+1}\in X} e^{r(u_i(x^t,\boldsymbol{a})+\beta v_i(x^{t+1}))} q(x^{t+1}|x^t,\boldsymbol{a})\, \boldsymbol{\varphi}(d\boldsymbol{a}|x^t). \end{aligned} \tag{3.4}$$

Let $T^{i,n}_{\boldsymbol{\varphi}}$ be the composition of $T^i_{\boldsymbol{\varphi}}$ with itself n times. Let $U_i^{[2,n]}(\boldsymbol{\varphi})(x^2)$ be the utility of player i in the $(n-1)$-stage subgame starting at state $x^2 \in X$.

Then from (3.4) it follows that

$$U_i^{[1,n]}(\boldsymbol{\varphi})(x^1) = T_{\boldsymbol{\varphi}}^{i,n}\mathbf{0}(x^1) = T_{\boldsymbol{\varphi}}^{i} U_i^{[2,n]}(\boldsymbol{\varphi})(x^1),$$
$$U_i(\boldsymbol{\varphi})(x^1) = T_{\boldsymbol{\varphi}}^{i} U_i(\boldsymbol{\varphi})(x^1). \tag{3.5}$$

Formula (3.5) says that we deal with *recursive utilities* satisfying Koopmans' equation [29] with $T_{\boldsymbol{\varphi}}^{i}$ as the aggregator of the utility obtained in the first stage and the utility from stage 2 onwards. Koopmans' approach [29] allows, however, for more general aggregators to study recursive utilities in infinite time horizon problems. Our aggregator is defined with the help of the entropic risk measure.

REMARK 3.6. Recursive utilities have attracted the interest of many researchers and found applications in stochastic decision models [3, 9, 10, 21, 36] and games [2, 4, 27]. However, the last two papers on games deal with intergenerational models, in which the solution is defined as a subgame perfect equilibrium. In [2], on the other hand, the authors examined games of resource extraction, in which the players have identical risk-sensitive preferences and apply discounted recursive utilities. Exploiting the form of a one-stage payoff and a transition probability they obtained a stationary non-randomised equilibrium in the class of non-decreasing right-continuous functions on the state space being a subset of the real numbers. The proof relies on the Schauder–Tikhonov fixed point theorem. Hence, the model and the tools used in [2] are different than in this work.

3.2. DSGs with the entropic risk measure on the space of all plays. In this section we describe an alternative approach and results for stochastic games on a countable state space obtained in [8], when the entropic risk measure is not used step by step but it is applied on the space of all plays.

Let $H^\infty = \mathbb{K} \times \mathbb{K} \times \cdots$ be the space of all infinite histories of the game (plays) endowed with the product σ-algebra. For any profile of strategies $\boldsymbol{\sigma} \in \Sigma$, a probability measure $\mathbb{P}_{x^1}^{\boldsymbol{\sigma}}$ and a stochastic process $(x^t, \boldsymbol{a}^t)_{t\in\mathbb{N}}$ are defined on H^∞ in a canonical way according to the Ionescu–Tulcea theorem [13, Proposition 7.28]. Here x^1 is an initial state. The expectation operator with respect to $\mathbb{P}_{x^1}^{\boldsymbol{\sigma}}$ is denoted by $\mathbb{E}_{x^1}^{\boldsymbol{\sigma}}$.

For each $i \in \mathcal{N}$, the *discounted utility function* defined involving the risk measure on the space H^∞ is defined as follows:

$$\widetilde{U}_i(\boldsymbol{\sigma})(x^1) := \frac{1}{r} \ln \mathbb{E}_{x^1}^{\boldsymbol{\sigma}} \left[e^{r \sum_{t=1}^{\infty} \beta^{t-1} u_i(x^t, \boldsymbol{a}^t)} \right]. \tag{3.6}$$

This form of utility was used in a number of papers on risk-sensitive control processes, e.g., [11, 16, 17], and on dynamic games, e.g., [7, 8, 5, 12, 15, 28]. The drawback of the model with the utility in (3.6) is that an optimal stationary policy/strategy need not exist. This is because the real discount factor between consecutive periods is not constant as in the risk-neutral case; see a detailed discussion in [16]. Therefore, one can expect to have optimality in the Markovian class of policies.

Basu and Ghosh [8] considered stochastic games with a countable state space and proved that any DSG with the utility functions in (3.6) has a Nash equilibrium $\boldsymbol{\sigma}^* = (\sigma_1^*, \ldots, \sigma_N^*)$ where every σ_i^* is Markovian, i.e., $\sigma_i^* = (\sigma_i^{*t})_{t\in\mathbb{N}}$ and $\sigma_i^{*t}(da_i \mid h^t) = \sigma_i^{*t}(da_i|x^t)$ for all $t \in \mathbb{N}$. Thus, the mixed action chosen by each player depends on time t and the state x^t. In stationary equilibrium, only the current state matters. The Bellman equations in the two aforementioned approaches, i.e. with the utilities in (3.3) and utilities in (3.6), are different.

Remark 3.7. A non-stationary Nash equilibrium obtained in [8] can be seen as stationary if we replace the state space X by $X \times \mathbb{N}$ and change the transition probability in a standard way. The new transition function is $\tilde{q}((y, n+1)|(x,n), \boldsymbol{a}) = q(y|x, \boldsymbol{a})$. However, for the finite set X we obtain a stationary Nash equilibrium profile, where strategies of the players depend on infinitely many states in $X \times \mathbb{N}$. If X is finite and we deal with DSG with recursive utilities of the players, then a stationary Nash equilibrium consists of strategies depending only on finitely many states. This fact seems important if we think about algorithms for finding Nash equilibria.

3.3. A game with payoffs induced by the Markowitz risk measure. We give a bimatrix game which can be seen as a stochastic game starting at some state and then moving immediately to an absorbing state. The example illustrates that the Nash equilibrium need not exist if the entropic risk measure is replaced by the Markowitz measure. By the *Markowitz risk measure* of a random variable Z we mean $\mathbb{E}Z + \frac{2}{r}\operatorname{Var} Z$ with $r < 0$. For convenience, we assume that $r = -2$. The random variables in our example are payoffs when the players use mixed strategies.

Example 3.8. We assume that the players play the following bimatrix game and the game moves with probability 1 to an absorbing state with zero rewards. Hence, in fact we deal with a game with just one step. The payoff matrices for players 1 and 2 are

$$P_1 = \begin{bmatrix} 2 & -2 \\ -1 & 1 \end{bmatrix} \quad \text{and} \quad P_2 = \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}.$$

Note that this game has no pure Nash equilibrium. A strategy for player 1 is $\mu = (a, 1-a)$ where $0 \le a \le 1$. A strategy for player 2 is $\sigma = (b, 1-b)$ with $0 \le b \le 1$. The expected payoffs under a pair (μ, σ) are

$$p_1(\mu, \sigma) = (3a-1)(2b-1) \quad \text{and} \quad p_2(\mu, \sigma) = -(3a-1)(2b-1).$$

The variances are the same for both players:

$$\mathrm{Var}(p_1(\mu,\sigma)) = \mathrm{Var}(p_2(\mu,\sigma)) = 3a + 1 - (3a-1)^2(2b-1)^2.$$

Thus, the payoffs under the Markowitz measure are

$$m_1(\mu,\sigma) := (3a-1)(2b-1) - 3a - 1 + (3a-1)^2(2b-1)^2, \tag{3.7}$$

$$m_2(\mu,\sigma) := -(3a-1)(2b-1) - 3a - 1 + (3a-1)^2(2b-1)^2. \tag{3.8}$$

We now show that this game has no Nash equilibrium. Suppose, on the contrary, that the game with payoff functions m_1 and m_2 has a Nash equilibrium (μ^*,σ^*) and $\sigma^* = (b^*, 1-b^*)$. If $2b^*-1 = 0$, then from (3.7), the best response of player 1 is to choose $a = 0$. If $2b^*-1 \neq 0$, then the best response of player 1 is to choose $a = 0$ or $a = 1$. That is because (3.7) yields $m_1(\mu,\sigma^*) = c_1a^2 + c_2a + c_3$ with some $c_2, c_3 \in \mathbb{R}$ and $c_1 > 0$. Thus, μ^* is a pure strategy. By (3.8), we have $m_2(\mu^*,\sigma) = d_1b^2 + d_2b + d_3$ with some $d_2, d_3 \in \mathbb{R}$ and $d_1 > 0$. Hence, the best response of player 2 to μ^* is to choose $b = 0$ or $b = 1$. Thus, σ^* is a pure strategy. Let (μ^*,σ^*) be a pure Nash equilibrium in the game with payoff functions m_1 and m_2. Observe that $m_i(\mu^*,\sigma^*) = p_i(\mu^*,\sigma^*)$ for $i = 1,2$. Therefore, (μ^*,σ^*) should be a pure Nash equilibrium in the game with payoff functions p_1 and p_2. But this game has no pure Nash equilibrium. Since $\ln(\cdot)$ is increasing and $1/r < 0$, for $r = -2$ any Nash equilibrium $(\widehat{\mu},\widehat{\sigma})$ in the game with payoffs induced by the entropic risk measure is a Nash equilibrium in the bimatrix game with payoff matrices

$$\widehat{P}_1 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = -\begin{bmatrix} e^{-4} & e^{4} \\ e^{2} & e^{-2} \end{bmatrix}, \qquad \widehat{P}_2 = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = -\begin{bmatrix} e^{4} & e^{-4} \\ e^{-2} & e^{2} \end{bmatrix}.$$

By the well-known method, one can derive $\widehat{\mu} = (\widehat{a}, 1-\widehat{a})$ and $\widehat{\sigma} = (\widehat{b}, 1-\widehat{b})$ where

$$\widehat{a} = \frac{b_{22}-b_{21}}{b_{11}-b_{12}+b_{22}-b_{21}} = \frac{e^6-e^2}{e^8+e^6-e^2-1} \approx 0.1173,$$

$$\widehat{b} = \frac{a_{22}-a_{12}}{a_{11}-a_{12}+a_{22}-a_{21}} = \frac{e^8-e^2}{e^8+e^6-e^2-1} \approx 0.8808.$$

The absolute risk aversion coefficient $|r| = 2$ is rather high. It is confirmed if we compare the equilibrium $(\widehat{\mu},\widehat{\sigma})$ with the unique Nash equilibrium $(\widetilde{\mu},\widetilde{\sigma})$ in the risk-neutral case where $\widetilde{\mu} = (\widetilde{a}, 1-\widetilde{a})$ with $\widetilde{a} = 1/3 > \widehat{a}$ and $\widetilde{\sigma} = (\widetilde{b}, 1-\widetilde{b})$ with $\widetilde{b} = 1/2 < \widehat{b}$.

4. Proof of the Equilibrium Theorem. First we state some auxiliary lemmas.

Lemma 4.1. *Assume that $f_m, f \in B(X)$ and $\lim_{m\to\infty} f_m(x) = f(x)$ for all $x \in X$. If Assumption* **A** *holds and $\boldsymbol{a}_m \to \boldsymbol{a} \in A(x)$ as $m \to \infty$, then*

$$\lim_{m\to\infty} \sum_{y\in X} f_m(y)q(y|x,\boldsymbol{a}_m) = \sum_{y\in X} f(y)q(y|x,\boldsymbol{a}) \quad \text{for all } x \in X.$$

Proof. If we show that $\boldsymbol{a} \to q(Z|x,\boldsymbol{a})$ is continuous for each $Z \subset X$, then the lemma follows from [35, Chapter 11, Proposition 18]. Let $Z = \{z_1, z_2, \ldots\}$ and $Z_n = \{z_1, \ldots, z_n\}$. Then $q(Z|x,\boldsymbol{a}) = \sup_n q(Z_n|x,\boldsymbol{a})$ is lower semicontinuous on $A(x)$. Similarly, $q(X \setminus Z|x,\boldsymbol{a})$ is lower semicontinuous on $A(x)$. Since $q(Z|x,\boldsymbol{a}) = 1 - q(X \setminus Z|x,\boldsymbol{a})$, $\boldsymbol{a} \mapsto q(Z|x,\boldsymbol{a})$ is continuous. ∎

For any $\boldsymbol{a} = (a_1, \ldots, a_N) \in A(x)$, we use the standard notation $\boldsymbol{a}_{-i}$ for the action profile of all players except player i, i.e., $\boldsymbol{a}_{-i} \in A_{-i}(x) := \prod_{j\neq i} A_j(x)$. We identify $(\boldsymbol{a}_{-i}, a_i)$ with $\boldsymbol{a}$. If $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_N) \in \Phi$ and $i \in \mathcal{N}$, then $\boldsymbol{\varphi}_{-i}(d\boldsymbol{a}_{-i}|x) \in \Pr(A_{-i}(x))$ is the product measure induced by all $\varphi_j(da_j|x)$ with $j \in \mathcal{N} \setminus \{i\}$. Let $\boldsymbol{\varphi} \in \Phi$, $i \in \mathcal{N}$ and $v \in B(X)$. For $x \in X$ we define

$$F^i_{\boldsymbol{\varphi}_{-i}} v(x) := \max_{\mu\in\Pr(A_i(x))} \mathcal{E}_i(\boldsymbol{\varphi}_{-i}, \mu, v)(x),$$

where

$$\mathcal{E}_i(\boldsymbol{\varphi}_{-i}, \mu, v)(x) := \frac{1}{r} \ln \int_{A_i(x)} \int_{A_{-i}(x)} \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a}_{-i},a_i)+\beta v(y))} q(y|x,\boldsymbol{a}_{-i},a_i)\, \boldsymbol{\varphi}_{-i}(d\boldsymbol{a}_{-i}|x)\, \mu(da_i).$$

Lemma 4.2. *$F^i_{\boldsymbol{\varphi}_{-i}} : B(X) \to B(X)$ is a contraction mapping for $i \in \mathcal{N}$.*

Proof. We apply the same arguments as Blackwell [14], who dealt with standard discounted dynamic programming. If $v, w \in B(X)$ and $v \le w$, then $F^i_{\boldsymbol{\varphi}_{-i}} v \le F^i_{\boldsymbol{\varphi}_{-i}} w$. Moreover, for any $v \in B(X)$ and $c \in \mathbb{R}$, $F^i_{\boldsymbol{\varphi}_{-i}}(v + c) = F^i_{\boldsymbol{\varphi}_{-i}} v + \beta c$. Therefore, for any $v, w \in B(X)$ we can write

$$F^i_{\boldsymbol{\varphi}_{-i}} v = F^i_{\boldsymbol{\varphi}_{-i}}(w + v - w) \le F^i_{\boldsymbol{\varphi}_{-i}}(w + \|v - w\|) \le F^i_{\boldsymbol{\varphi}_{-i}} w + \beta\|v - w\|.$$

Hence, $F^i_{\boldsymbol{\varphi}_{-i}} v - F^i_{\boldsymbol{\varphi}_{-i}} w \le \beta\|v - w\|$. Similarly, we get $F^i_{\boldsymbol{\varphi}_{-i}} w - F^i_{\boldsymbol{\varphi}_{-i}} v \le \beta\|v - w\|$, which finishes the proof. ∎

Assume that $\boldsymbol{\phi} \in \Phi$, $i \in \mathcal{N}$ and $v \in B(X)$. Recall that

$$T^i_{\boldsymbol{\phi}} v(x) = \frac{1}{r} \ln \int_{A(x)} \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta v(y))} q(y|x,\boldsymbol{a})\, \boldsymbol{\phi}(d\boldsymbol{a}|x).$$

Lemma 4.3.

(a) *$T^i_{\boldsymbol{\phi}} : B(X) \to B(X)$ is a contraction mapping.*
(b) *$U_i(\boldsymbol{\phi})$ is the unique fixed point of $T^i_{\boldsymbol{\phi}}$ and $\lim_{n\to\infty} \|U_i(\boldsymbol{\phi}) - T^{i,n}_{\boldsymbol{\phi}} v\| = 0$ for any $v \in B(X)$.*

(c) $\|U_i(\boldsymbol{\phi}) - U_i^{[1,n]}(\boldsymbol{\phi})\| \le \frac{L\beta^n}{1-\beta}$ *for* $L > 0$ *such that* $0 \le u_i \le L$ *and for all* $\boldsymbol{\phi} \in \Phi$ *and* $n \in \mathbb{N}$.

Proof. Part (a) is obvious. Parts (b) and (c) are corollaries to the Banach contraction mapping theorem. In particular,

$$\|U_i(\boldsymbol{\phi}) - U_i^{[1,n]}(\boldsymbol{\phi})\| = \|U_i(\boldsymbol{\phi}) - T_{\boldsymbol{\phi}}^{i,n}\mathbf{0}\| \le \frac{\beta^n}{1-\beta}\|T_{\boldsymbol{\phi}}^{i}\mathbf{0} - \mathbf{0}\| \le \frac{L\beta^n}{1-\beta}. \blacksquare$$

For $i \in \mathcal{N}$ let $C_i^* := \prod_{x\in X} C^*(A_i(x))$ be the countable product of the topological dual spaces $C^*(A_i(x))$ of the Banach spaces $C(A_i(x))$. Assume that C_i^* is endowed with the product of the weak-star topologies on $C^*(A_i(x))$. Since all action spaces $A_i(x)$ are compact metric, C_i^* is metrisable. It is also a locally convex linear topological space (see [1, Chapter 5.14]). Note that by the Riesz representation theorem, every $\phi_i = (\phi_i(\cdot|x))_{x\in X} \in \Phi_i$ can be recognised as an element of a compact convex set $\prod_{x\in X} \Pr(A_i(x)) \subset C_i^*$. Compactness of the product space follows from Tikhonov's theorem. Therefore, we can think that Φ_i is a compact convex subset of a locally convex metrisable vector space. The convergence $\phi_{i,n} \to \phi_{i,o}$ in Φ_i means that $\phi_{i,n}(\cdot|x) \to^* \phi_{i,o}(\cdot|x)$ as $n \to \infty$ for all $x \in X$. Assume that $\Phi = \prod_{i\in\mathcal{N}} \Phi_i$ is given the product topology. Let $\boldsymbol{\phi}_n = (\phi_{1,n}, \ldots, \phi_{N,n})$ and $\boldsymbol{\phi}_o = (\phi_{1,o}, \ldots, \phi_{N,o})$ belong to Φ. The following fact is important for our proof (see [13, Chapter 7.4]). If $\boldsymbol{\phi}_n \to \boldsymbol{\phi}_o$ in Φ and $g \in C(A_1(x) \times \cdots \times A_N(x))$ for every $x \in X$, then

$$\lim_{n\to\infty} \int_{A(x)} g(\boldsymbol{a})\,\boldsymbol{\phi}_n(d\boldsymbol{a}|x) = \int_{A(x)} g(\boldsymbol{a})\,\boldsymbol{\phi}_o(d\boldsymbol{a}|x) \quad \text{for all } x \in X. \tag{4.1}$$

Here we have $\boldsymbol{\phi}_k(d\boldsymbol{a}|x) = \phi_{1,k}(da_1|x) \otimes \cdots \otimes \phi_{N,k}(da_N|x)$, $k = o$ or $k = n$.

LEMMA 4.4. *Under Assumption* **A**, *for each* $i \in \mathcal{N}$ *and* $x \in X$, *the function* $\boldsymbol{\phi} \mapsto U_i(\boldsymbol{\phi})(x)$ *is continuous on* Φ.

Proof. From Lemma 4.3(c), it follows that it is sufficient to prove the assertion for all n-stage games with $n \in \mathbb{N}$. For 1-stage games, it is a simple corollary to (4.1). Now we prove the induction step. Assume that $\boldsymbol{\phi}_k \to \boldsymbol{\phi}_o$ in Φ and

$$g_k(y) := U_i^{[2,n]}(\boldsymbol{\phi}_k)(y) \to g_o(y) := U_i^{[2,n]}(\boldsymbol{\phi}_o)(y) \quad \text{as } k \to \infty.$$

Here we recall that $U_i^{[2,n]}(\boldsymbol{\phi})(y)$ denotes the utility in the $(n-1)$-stage game from period 2 to n, if the profile $\boldsymbol{\phi}$ of stationary strategies is used and $x^2 = y$. Note that since $r < 0$ and $u_i \ge 0$, we have

$$\begin{aligned} &\Big|\sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta g_k(y))} q(y|x,\boldsymbol{a}) - \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta g_o(y))} q(y|x,\boldsymbol{a})\Big| \\ &\qquad \le z_k(x) := \sup_{\boldsymbol{a}\in A(x)} \sum_{y\in X} |e^{r\beta g_k(y)} - e^{r\beta g_o(y)}| q(y|x,\boldsymbol{a}). \end{aligned} \tag{4.2}$$

We wish to show that $\lim_{k\to\infty} z_k(x) = 0$. Suppose, on the contrary, that there exist $\epsilon > 0$ and an infinite set $\mathbb{N}_1 \subset \mathbb{N}$ such that $z_{k'}(x) \geq \epsilon$ for all $k' \in \mathbb{N}_1$. Under Assumption **A**, from Lemma 4.1 it follows that for each $k' \in \mathbb{N}_1$ there exists $\boldsymbol{a}_{k'} \in A(x)$ such that

$$z_{k'}(x) = \sum_{y\in X} |e^{r\beta g_{k'}(y)} - e^{r\beta g_o(y)}| q(y|x, \boldsymbol{a}_{k'}).$$

Without loss of generality, assume that $\boldsymbol{a}_{k'} \to \boldsymbol{a}' \in A(x)$ as $k' \to \infty$. Using Lemma 4.1 again, we find that $\lim_{k'\to\infty} z_{k'}(x) = 0$. This contradiction proves that $\lim_{k\to\infty} z_k(x) = 0$. Now from (4.2), it follows that

$$\sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta g_k(y))} q(y|x, \boldsymbol{a}) \to \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta g_o(y))} q(y|x, \boldsymbol{a})$$

as $k \to \infty$, uniformly in $\boldsymbol{a} \in A(x)$. This fact and the convergence of $\boldsymbol{\phi}_k(\cdot|x)$ to $\boldsymbol{\phi}_o(\cdot|x)$ in $\Pr(A(x))$ imply that

$$\int_{A(x)} \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta g_k(y))} q(y|x, \boldsymbol{a})\, \boldsymbol{\phi}_k(d\boldsymbol{a}|x)$$
$$\to \int_{A(x)} \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a})+\beta g_o(y))} q(y|x, \boldsymbol{a})|\, \boldsymbol{\phi}_o(d\boldsymbol{a}|x) \quad \text{as } k \to \infty.$$

Finally, the above convergence leads to

$$\lim_{k\to\infty} U_i^{[1,n]}(\boldsymbol{\phi}_k)(x) = U_i^{[1,n]}(\boldsymbol{\phi}_o)(x).$$

This completes the induction step. ■

Assume that $\boldsymbol{\varphi} \in \Phi$ and $\sigma_i = (\sigma_i^t)_{t\in\mathbb{N}} \in \Sigma_i$. For the proof of the Equilibrium Theorem we need the following operator:

$$T^i_{[\boldsymbol{\varphi}_{-i},\sigma_i^t]} v_i(h^t) = \frac{1}{r} \ln\Big(\int_{A_{-i}(x^t)} \int_{A_i(x^t)} \sum_{x^{t+1}\in X} e^{r(u_i(x^t,\boldsymbol{a}_{-i},a_i)+\beta v_i(h^t,\boldsymbol{a}_{-i},a_i,x^{t+1}))}$$
$$\times\, q(x^{t+1}|x^t, \boldsymbol{a}_{-i}, a_i) \boldsymbol{\varphi}_{-i}(d\boldsymbol{a}_{-i}|x^t) \sigma_i^t(da_i|h^t)\Big).$$

This operator is monotone, i.e.,

$$T^i_{[\boldsymbol{\varphi}_{-i},\sigma_i^t]} v_i(h^t) \geq T^i_{[\boldsymbol{\varphi}_{-i},\sigma_i^t]} v_i'(h^t) \quad \text{if} \quad v_i(h^{t+1}) \geq v_i'(h^{t+1}) \text{ for all } h^{t+1} \in H^{t+1}.$$

Proof of the Equilibrium Theorem. Let $\boldsymbol{\varphi} \in \Phi$. For each player $i \in \mathcal{N}$ we define

$$\mathcal{BR}_i(\boldsymbol{\varphi}_{-i}) = \big\{\psi_i \in \Phi_i : U_i([\boldsymbol{\varphi}_{-i}, \psi_i])(x) \geq U_i([\boldsymbol{\varphi}_{-i}, \phi_i])(x) \text{ for all } \phi_i \in \Phi_i,\ x \in X\big\}.$$

The best response correspondence is defined as

$$\boldsymbol{\varphi} \mapsto \mathcal{BR}(\boldsymbol{\varphi}) := \prod_{i\in\mathcal{N}} \mathcal{BR}_i(\boldsymbol{\varphi}_{-i}).$$

Put $w_i(x) := U_i([\boldsymbol{\varphi}_{-i}, \psi_i])(x)$ for $x \in X$. Notice that $\psi_i \in \mathcal{BR}_i(\boldsymbol{\varphi}_{-i})$ if and only if

$$w_i(x) = F^i_{\boldsymbol{\varphi}_{-i}} w_i(x) = \mathcal{E}_i(\boldsymbol{\varphi}_{-i}, \psi_i(\cdot|x), w_i)(x),$$

and, since $r < 0$, this is equivalent to the statement that $\psi_i(\cdot|x)$ satisfies the equality

$$\int\limits_{A_i(x)} \int\limits_{A_{-i}(x)} \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a}_{-i},a_i)+\beta w_i(y))} q(y|x,\boldsymbol{a}_{-i},a_i)\,\boldsymbol{\varphi}_{-i}(d\boldsymbol{a}_{-i}|x)\psi_i(da_i|x)$$
$$= \min_{\mu\in\Pr(A_i(x))} \int\limits_{A_i(x)} \int\limits_{A_{-i}(x)} \sum_{y\in X} e^{r(u_i(x,\boldsymbol{a}_{-i},a_i)+\beta w_i(y))}$$
$$\times\, q(y|x,\boldsymbol{a}_{-i},a_i)\,\boldsymbol{\varphi}_{-i}(d\boldsymbol{a}_{-i}|x)\mu(da_i).$$

Hence, $\mathcal{BR}_i(\boldsymbol{\varphi}_{-i})$ is non-empty and convex. From the continuity of the utility functions U_i (Lemma 4.4), it follows that the correspondence $\boldsymbol{\varphi} \mapsto \mathcal{BR}(\boldsymbol{\varphi})$ is upper semicontinuous. By the Kakutani–Fan–Glicksberg fixed point theorem [1, Corollary 17.55], there exists $\boldsymbol{\varphi}^* \in \Phi$ such that $\boldsymbol{\varphi}^* \in \mathcal{BR}(\boldsymbol{\varphi}^*)$. From the definition of $\mathcal{BR}(\boldsymbol{\varphi}^*)$, it follows that $\boldsymbol{\varphi}^*$ is a Nash equilibrium in the class of all stationary strategy profiles.

Next we show that $\boldsymbol{\varphi}^*$ is a Nash equilibrium in the class of all strategy profiles. Fix any player $i \in \mathcal{N}$ and consider any $\sigma_i = (\sigma_i^t)_{t\in\mathbb{N}} \in \Sigma_i$. Put $w_i^*(x) := U_i(\boldsymbol{\varphi}^*)(x)$ for $x \in X$. We have $w_i^*(x) = F^i_{\boldsymbol{\varphi}^*_{-i}} w_i^*(x)$ for each $x \in X$. This implies that, for all $t \in \mathbb{N}$, $x^t \in X$ and all $h^t \in H^t$,

$$w_i^*(x^t) = F^i_{\boldsymbol{\varphi}^*_{-i}} w_i^*(x^t) \ge T^i_{[\boldsymbol{\varphi}^*_{-i},\sigma_i^t]} w_i^*(h^t). \tag{4.3}$$

From (4.3) and the monotonicity of the operator $T^i_{[\boldsymbol{\varphi}^*_{-i},\sigma_i^t]}$ we conclude that

$$\begin{aligned} w_i^*(x^1) &\ge T^i_{[\boldsymbol{\varphi}^*_{-i},\sigma_i^1]} \cdots T^i_{[\boldsymbol{\varphi}^*_{-i},\sigma_i^n]} w_i^*(x^1) \\ &\ge T^i_{[\boldsymbol{\varphi}^*_{-i},\sigma_i^1]} \cdots T^i_{[\boldsymbol{\varphi}^*_{-i},\sigma_i^n]} \mathbf{0}(x^1) = U_i^{[1,n]}([\boldsymbol{\varphi}^*_{-i},\sigma_i])(x^1) \end{aligned}$$

for all $x^1 \in X$. Hence,

$$w_i^*(x^1) = U_i(\boldsymbol{\varphi}^*)(x^1) \ge \lim_{n\to\infty} U_i^{[1,n]}([\boldsymbol{\varphi}^*_{-i},\sigma_i])(x^1) = U_i([\boldsymbol{\varphi}^*_{-i},\sigma_i])(x^1)$$

for all $x^1 \in X$, which completes the proof. ∎

Acknowledgments. This paper is supported by NCN Grant No. UMO-2022/47/B/HS4/00331.

References

[1] C. Aliprantis and K. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, New York, 2006.

[2] H. Asienkiewicz and Ł. Balbus, *Existence of Nash equilibria in stochastic games of resource extraction with risk-sensitive players*, TOP 27 (2019), 502–518.
[3] H. Asienkiewicz and A. Jaśkiewicz, *A note on a new class of recursive utilities in Markov decision processes*, Appl. Math. (Warsaw) 44 (2017), 149–161.
[4] Ł. Balbus, *Markov perfect equilibria in OLG models with risk sensitive agents*, Topol. Methods Nonlinear Anal. 53 (2019), 309–333.
[5] T. Başar, *Nash equilibria of risk-sensitive nonlinear stochastic differential games*, J. Optim. Theory Appl. 100 (1999), 479–498.
[6] T. Başar, *Robust designs through risk sensitivity: an overview*, J. Systems Sci. Complexity 34 (2021), 1634–1665.
[7] A. Basu and M. K. Ghosh, *Zero-sum risk-sensitive stochastic games on a countable state space*, Stoch. Process. Appl. 124 (2014), 961–983.
[8] A. Basu and M. K. Ghosh, *Nonzero-sum risk-sensitive stochastic games on a countable state space*, Math. Oper. Res. 43 (2018), 516–532.
[9] N. Bäuerle and A. Jaśkiewicz, *Optimal dividend payout model with risk sensitive preferences*, Insurance Math. Econom. 73 (2017), 82–93.
[10] N. Bäuerle and A. Jaśkiewicz, *Stochastic optimal growth model with risk sensitive preferences*, J. Econom. Theory 173 (2018), 181–200.
[11] N. Bäuerle and U. Rieder, *More risk-sensitive Markov decision processes*, Math. Oper. Res. 39 (2014), 105–120.
[12] N. Bäuerle and U. Rieder, *Zero-sum risk-sensitive stochastic games*, Stoch. Process. Appl. 27 (2017), 622–642.
[13] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: the Discrete-Time Case*, Academic Press, New York, 1978.
[14] D. Blackwell, *Discounted dynamic programming*, Ann. Math. Statist. 36 (1965), 226–235.
[15] P. Caravani and G. Papavassilopoulos, *A class of risk-sensitive noncooperative games*, J. Econom. Dynam. Control 14 (1990), 117–149.
[16] K.-J. Chung and M. Sobel, *Discounted MDP's: distribution functions and exponential utility maximization*, SIAM J. Control Optim. 25 (1987), 49–62.
[17] G. B. Di Masi and Ł. Stettner, *Infinite horizon risk sensitive control of discrete time Markov processes with small risk*, Systems Control Lett. 40 (2000), 15–20.
[18] A. Federgruen, *On N-person stochastic games with denumerable state space*, Adv. Appl. Probab. 10 (1978), 452–471.
[19] A. M. Fink, *Equilibrium in a stochastic n-person game*, J. Sci. Hiroshima Univ. Ser. A-I Math. 28 (1964), 89–93.
[20] H. Föllmer and A. Schied, *Stochastic Finance*, de Gruyter, Berlin, 2004.
[21] L. P. Hansen and T. J. Sargent, *Discounted linear exponential quadratic Gaussian control*, IEEE Trans. Autom. Control 40 (1995), 968–971.
[22] R. A. Howard and J. E. Matheson, *Risk-sensitive Markov decision processes*, Management Sci. 18 (1972), 356–369.
[23] A. Idzik, *Two-stage noncooperative stochastic games with denumerable state spaces*, in: Optimization Techniques, Lecture Notes in Control Information Sci. 22, Springer, Berlin, 1980.
[24] D. Jacobson, *Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games*, IEEE Trans. Autom. Control 18 (1973), 124–131.
[25] S. C. Jaquette, *Markov decision processes with a new optimality criterion: discrete time*, Ann. Statist. 1 (1973), 496–505.

[26] S. C. Jaquette, *A utility criterion for Markov decision processes*, Management Sci. 23 (1976), 43–49.
[27] A. Jaśkiewicz and A. S. Nowak, *Stationary Markov perfect equilibria in risk sensitive stochastic overlapping generations models*, J. Econom. Theory 151, (2014), 411–447.
[28] M. B. Klompstra, *Nash equilibria in risk-sensitive dynamic games*, IEEE Trans. Autom. Control 45 (2000), 1397–1401.
[29] T. C. Koopmans, *Stationary ordinal utility and impatience*, Econometrica 28 (1960), 287–309.
[30] D. G. Luenberger, *Investment Science*, 2nd ed., Oxford Univ. Press, 2014.
[31] J. Miao, *Economic Dynamics in Discrete Time*, MIT Press, Cambridge, MA, 2014.
[32] G. E. Monahan and M. J. Sobel, *Risk-sensitive dynamic Market share attraction games*, Games Econom. Behavior 20 (1997), 149–160.
[33] M. Pitera and Ł. Stettner, *Discrete-time risk sensitive portfolio optimization with proportional transaction costs*, Math. Finnance 33 (2023), 1287–1313.
[34] J. W. Pratt, *Risk aversion in the small and in the large*, Econometrica 32 (1964), 122–136.
[35] H. Royden, *Real Analysis*, Macmillan, New York, 1968.
[36] T. J. Sargent and J. Stachurski, *Dynamic Programming. Vol. I*, Foundations Quantecon, 2023.
[37] M. Takahashi, *Equilibrium points of stochastic non-cooperative n-person games*, J. Sci. Hiroshima Univ. Ser A-I, Math. 28 (1964), 95–99.
[38] H. Tembine, Q. Zhu and T. Başar, *Risk-sensitive mean-field games*, IEEE Trans. Autom. Control 59 (2014), 835–850.
[39] P. Weil, *Precautionary savings and the permanent income hypothesis*, Rev. Econom. Stud. 60 (1993), 367–383.
[40] P. Whittle, *Risk-sensitive linear/quadratic/Gaussian control*, Adv. Appl. Probab. 13 (1981), 764–777.
[41] P. Whittle, *Risk-Sensitive Optimal Control*, Wiley, New York, 1990.

Łukasz Balbus, Andrzej S. Nowak
Faculty of Mathematics, Computer Science and Econometrics
University of Zielona Góra
Zielona Góra, Poland
E-mail: l.balbus@wmie.uz.zgora.pl
a.nowak@wmie.uz.zgora.pl
ORCID: 0000-0002-3181-5187

Anna Jaśkiewicz
Faculty of Pure and Applied Mathematics
Wrocław University of Science and Technology
Wrocław, Poland
E-mail: anna.jaskiewicz@pwr.edu.pl