Hindawi
Advances in Operations Research
Volume 2022, Article ID 3184610, 13 pages
https://doi.org/10.1155/2022/3184610

Research Article

A Consumption and Investment Problem via a Markov Decision Processes Approach with Random Horizon

Octavio Paredes Pérez, Víctor Hugo Vázquez Guevara, and Hugo Cruz-Suárez

Facultad de Ciencias Físico Matemáticas, Benemérita Universidad Autónoma de Puebla, Avenida San Claudio y Río Verde, San Manuel 72570, Puebla, Mexico

Correspondence should be addressed to Víctor Hugo Vázquez Guevara; vvazquez@fcfm.buap.mx

Received 18 November 2021; Revised 5 April 2022; Accepted 13 May 2022; Published 7 June 2022

Academic Editor: Bikash Koli Dey

This work is devoted to a consumption and investment problem, in which there is an investor with certain initial wealth with the possibility of deciding how much of such wealth will be consumed and how much will be invested in each of a series of successive times. The key issue is to find a wealth assignation rule in order to maximize the performance criteria; such dilemma will be achieved by the dynamic programming technique for the Markov decision processes with random horizon.

1. Introduction

Markov decision processes (MDPs) provide a very useful system for creating and implementing a decision-making process whose results are partially random. MDPs are useful stochastic processes for boarding a wide range of optimization problems of continuous or discrete nature (In this paper, it will be only considered the discrete framework). In all the sequel, at each step, the process is in some state and the decision maker may choose any action that is available for such a state. The process responds at the next stage by randomly moving to a new state and giving a reward to the decision maker. The central problem of MDPs is to find an "optimal policy"; i.e., a function that specifies some mechanism for selecting actions optimally at each stage.

MDPs can be solved by dynamic programming. For example, in [1], a comprehensive and theoretical treatment of the mathematical foundations of optimal stochastic control of discrete-time systems is given; meanwhile, in [2], interest is mostly limited to MDPs with a Borel space of states and possibly unlimited costs. In [3], it is explained that the theory of the stochastic dynamic programming method is easily applicable to many practical problems, even for nonstationary models.

However, there exist another method which may be considered for solving stochastic optimization problems, for example: In [4], an emended minimax method is developed based on the semi-autonomized multiobjective optimization algorithm by amending the classical minimax method, which leads to desirable optimal values in the certitude state and to find another Pareto optimal solution under fuzziness in the incertitude state. In [5], the author focuses on one general fractile criterion iterative-interactive optimization process in order to obtain the preferable Pareto optimal solution, subject to a specified main objective function to multiobjective stochastic linear programming problems in a fuzzy environment. In addition, in [6], one real-life-based cost-effective and customer-centric closed-loop supply chain management model is considered together with the T set that represents the inherent impreciseness to objective functions which conducts to find that optimal values are superior than stipulated goals to both the objective functions in the T environment. In [7], the effects of setup cost reduction and quality improvement in a two-echelon supply chain model with deterioration are developed. The objective is to minimize the total cost of the entire supply chain model by simultaneously optimizing setup cost, process quality, number of deliveries, and lot size. In [8], a set of very interesting situations coming from mobile and wireless networks, connection management, and Internet is considered in which optimal decisions are required and it is necessary to provide a side view about control problems and the theory behind them.

In this paper, the possibility will be taken into account that external factors may force the process to be completed earlier than planned. In this way, it is necessary to consider the horizon as a random variable, which may be independent of the state-action space [9]. Such an idea has been explored; for example, in [10], where the optimal selection strategy for the Armed Bandit paradigm with random horizon and possibly random discount factors is found.

Hence, it will be considered and investor with certain initial capital which; in each of a random number of times, may reinvest into risky assets, consume or invest in a risk-free bond. The goal is to conceive a strategy of consumption and investment in order to maximize the expected sum of an utility coming from; exclusively, the spent capital at each stage. Hence, in this paper, via the theory of MDPs with a finite random horizon, an optimal policy of consumption and investment will be established, in the case in which utility function responsible for evaluating consumption is of the exponential type. Although these kind of utility functions are rather classical, they are as well useful since such functions consider a constant absolute aversion to risk and are the only risk-averse increasing utility function whose risk premium is invariant with respect to wealth [11, 12]. The term risk aversion refers to the preference for stochastic realizations with limited deviation from the expected value. In risk-averse optimal control, one may prefer a policy with a higher cost in expectation but lower deviations than one with a lower cost but possibly higher deviations [13].

This work is organized as follows: in section two, fundamental ideas around MDP with random horizon jointly with an equivalence between performance criteria of an MDP with a finite random horizon and the one associated with an MDP with deterministic horizon are analyzed. Section three will be addressed to a consumption and investment scenario with random horizon together with a numerical experiment. Additionally, this section contains the main contribution of this text: the finding of the optimal policy for the consumption and investment problem with a finite random horizon and exponential utility function. Finally, two appendices are included. On the one hand, Appendix A deals with basic definitions around the MDP theory together with some useful assumptions for solving the consumption and investment problem by means of the dynamic programming technique. On the other hand, the concept of the financial market will be discussed in Appendix B.

2. Markov Decision Process with a Random Horizon

In literature, one may find references where discrete-time control problems with a random horizon are discussed, for example: [9, 10]. Therefore, let τ be a discrete random variable associated with some probability space $(\Omega', \mathscr{F}', \mathbb{P})$. Suppose that the mass function of τ is known and given by $\rho_n := \mathbb{P}(\tau = n)$, with $n = 0, 1, 2, \ldots, N$, where N is a natural number or $N = \infty$. Consider now a Markov decision model (E, A, D, Q, r_n, g) and define the following performance criteria:

$$V^{\tau}(\pi, x) := \mathbb{E}\left[\sum_{n=0}^{\tau-1} r_n(X_n, a_n) + g(X_\tau)\right], \tag{1}$$

where $\pi \in \Delta$, $x \in E$, and $\mathbb{E}$ denotes the expected value with respect to the joint distribution of the process $\{(x_n, a_n)\}$ and τ. In order to introduce the corresponding optimal control problem, we define the optimal value function as follows:

$$V^{\tau}(x) := \sup_{\pi\in\Delta} V^{\tau}(\pi, x), \quad x \in E. \tag{2}$$

In this way, the optimal control problem with a random horizon consists of finding a policy $\pi^* \in \Delta$ such that $V^{\tau}(\pi^*, x) = V^{\tau}(x)$, $\forall x \in E$. The following assumption will be considered for simplifying the performance criteria under a discrete random horizon [9].

Assumption 1. For each $x \in E$ and $\pi \in \Delta$ induced process $\{(X_n, a_n) | n = 0, 1, 2, \ldots\}$ is independent of τ.

So, under Assumption 1 and equation (1), reward performance criteria are reduced to [9].

$$V^{\tau}(\pi, x) = \mathbb{E}_x^{\pi}\left[\sum_{n=0}^{N-1} \mathbb{P}_n r(X_n, a_n) + \mathbb{P}_N g(X_N)\right], \quad x \in E, \pi \in \Delta, \tag{3}$$

where for each $n = 0, 1, \ldots, N$, $\mathbb{P}_n := \mathbb{P}[\tau \geq n]$.

Thus, the optimal control problem with a random horizon τ is equivalent to the optimal control problem with planning horizon $N + 1$, a nonhomogeneous reward function and an equally zero terminal reward. Hence, Theorem A.6 may be considered under conditions of Assumption A.5. An alternative approach discussed in [9] considers a different set of assumption on the reward function (which remain fixed at each stage) and the transition kernel Q.

3. Consumption and Investment Problem with a Random Horizon

An investor has an initial wealth of $x > 0$ and at the beginning of at most N periods (implicitly a random horizon τ with support on $\{0, 1, \ldots, N\}$ is contemplated) he/she can determine which part of his/her wealth will be consumed and which part will be invested in the financial market given in Appendix B. Amount c_n denotes the consumed amount at time n and will be evaluated by an utility function U_c. Remaining money will be invested in d risky assets and in a risk-free bond. Terminal riches (X_τ) is judged via another utility function U_p. The main problem is designing a strategy of sequential decisions of consumption and investment in order to maximize the sum of his/her expected gains.

In all the sequel, Assumption B.9 will be supposed and no arbitrage opportunities are available. In addition, it is supposed that the domain of both utility functions U_c and

U_p is $[0, \infty)$. In this context, dynamics of the wealth is as follows:

$$X_{n+1} = (1 + i_{n+1})(X_n - c_n + a_n \cdot R_{n+1}), \quad (4)$$

where (c_n, a_n) is a consumption and investment strategy, i.e., (a_n) and (c_n) are $(\mathcal{F}_n)$ adapted, $0 \le c_n \le X_n$ and $X_0 = x$.

The consumption and investment problem previously described may be associated with a Markov decision model with the following components:

$E := [0, \infty)$,

$A := \mathbb{R}^+ \times \mathbb{R}^d$,

$D(x) := \{(c, a) \in A | 0 \le c \le x \text{ and } (1 + i_{n+1})(x - c + a \cdot R_{n+1}) \ge 0, \mathbb{P} - a.s.\}$ for $x \ge 0$,

$\mathcal{Z} := [-1, \infty)^d$ where $z \in \mathcal{Z}$ denotes the relative risk,

$T(x, c, a, z) := (1 + i_{n+1})(x - c + a \cdot z)$ is the transition function,

Q^Z is the distribution of R_{n+1},

$r(x, c, a) := U_c(c)$ is the reward function,

$g(x) := U_p(x)$ is the terminal reward function,

τ is the random horizon with support on $\{0, 1, \ldots, N\}$.

In this framework, value function is defined as follows:

$$V(x) = \sup_\pi \mathbb{E}_x^\pi \left[\sum_{n=0}^{\tau-1} U_c(c_n(X_n)) + U_p(X_\tau) \right], \quad (5)$$

where the supremum is taken over all policies $\pi = (f_0, \ldots, f_{N-1})$ with $f_n(x) = (c_n(x), a_n(x))$.

Sufficient conditions will be given to propose the solution of the consumption and investment problem with a random horizon with a finite support. Under assumption 1, the proof of the following result can be obtained by using Theorem A.6 and the wealth dynamics given in (4) [14]. Its conclusion allows to associate a MDP with a random horizon with support on $\{0, 1, \ldots, N\}$ with another MDP with a nonhomogenous reward function, deterministic horizon $(N + 1)$, and equally zero terminal cost.

Theorem 1. *In the multiperiodic consumption and investment problem, we define functions $V_0, V_1, \ldots, V_{N+1}$ on E by the following:*

$$V_{N+1}(x) := 0,$$
$$V_N(x) := \mathbb{P}_N U_p(x),$$
$$V_N(x) := \sup_{(c,a)\in D_n(x)} \cdot \{\mathbb{P}_n U_c(c) + \mathbb{E} V_{n+1}(1 + i_{n+1})(x - c + a \cdot R_{n+1})\}. \quad (6)$$

Then, there exist maximizers f_n^* of V_n and the strategy $(f_0^*, \ldots, f_N^*)$ is optimal for the consumption and investment problem.

3.1. Exponential Utility Function. This section deals with a version of the consumption and investment problem with exponential utility functions and a finite random horizon. In this setting, the process that describes the evolution of investor's capital may end before some fixed horizon due to external causes. However, Assumption 1 prevents the decision maker to finish such process because of bad investments, which may lead to drop this process below zero.

Utility functions arise naturally in economics and finance, for example:

On the mean-variance approach of Merton and Samuelson, it has already found that a quadratic utility provides a closed-form solution for the portfolio selection under very general conditions; however, on the case of power and the exponential utility function, there is no possibility to find closed-form solutions without information on the distribution of the return process [15]. In addition, in [16], by assuming that a portfolio's returns follow an approximate log-normal distribution, the closed-form expressions of the optimal portfolio weights were obtained for both power and logarithmic utility functions.

In portfolio optimization, in order to maximize the widespread logarithmic utility of some investor, assets whose prices depend on their past values in a non-Markovian way are taken into account [17]. On the same topic [18], Chapter 9 provides a very interesting contribution on the treatment of utility functions, in particular the risk aversion is deeply addressed.

On a similar matter, in [19], it is possible to review a self-contained survey of utility functions (exponential and power utilities of the first and second kind) together with some of their applications in finance. This reference also discusses the Pareto optimal risk exchanges and presents very illustrative examples dealing with earlier mentioned utility functions.

Exponential utility functions are widely employed because they consider a constant absolute aversion to risk [20–22], and they are the only risk-averse increasing utility functions whose risk premium is invariant respect to wealth [11, 12]. The term risk aversion refers to the preference for stochastic realizations with limited deviation from the expected value. In risk-averse optimal control, one may prefer a policy with a higher cost in expectation but lower deviations than one with lower cost but possibly higher deviations [13]. In addition, from the technical point of view, if both utility functions U_c and U_p are of the following form:

$$\left(-\frac{1}{\gamma}\right)\exp(-\gamma y), \quad \gamma > 0 \text{ and } y \in E, \quad (7)$$

then they are bounded superiorly, and hence, Assumption B.9 is directly satisfied.

The following assumption is needed by ensuring that optimal consumptions do not exceed available capital.

Assumption 2. For each $n = 0, 1, \ldots, N$ and $x > 0$, suppose that

$$1 < \frac{\mathbb{P}_n}{k_{n+1}\left(b_{n+1}\left(1 + i_{n+1}\right)v_n\right)} < \exp(x\gamma), \tag{8}$$

where

$$k_{n+1} = \left(\frac{\mathbb{P}_{n+1}}{\gamma}\right)\left(\frac{k_{n+2}b_{n+2}\left(1 + i_{n+2}\right)v_{n+1}}{\mathbb{P}_{n+1}}\right)^{\left(\gamma/b_{n+2}\left(1+i_{n+2}\right)+\gamma\right)} + k_{n+2}v_{n+1}\left(\frac{k_{n+2}b_{n+2}\left(1 + i_{n+2}\right)v_{n+1}}{\mathbb{P}_{n+1}}\right)^{\left(b_{n+2}\left(1+i_{n+2}\right)/\left(b_{n+2}\left(1+i_{n+2}\right)+\gamma\right)\right)}, \tag{9}$$

$$b_{n+1} = \frac{b_{n+2}\left(1 + i_{n+2}\right)\gamma}{b_{n+2}\left(1 + i_{n+2}\right) + \gamma}, \tag{10}$$

with $k_N = \mathbb{P}_N(1/\gamma)$, $b_N = \gamma$, $\mathbb{P}_n := \sum_{i=n}^{N}\rho_i = \mathbb{P}(\tau \geq n)$, $n = 0, 1, 2, \ldots, N$ and

$$v_n := \inf_{\alpha\in\mathbb{R}^d}\mathbb{E}\left[\exp\left(-\gamma\frac{S_N^0}{S_n^0}\alpha \cdot R_{n+1}\right)\right], \tag{11}$$

which exists, thanks to Theorem 4.1.1 of [14].

Then, it is possible to deduce the following result for the consumption and investment problem with a (finite) random horizon and risk-averse increasing utility function whose risk premium is invariant with respect to wealth (exponential utility function). This theorem provides a mechanism for acting optimally with respect to the consumption and risk investments at each stage, such optimal decisions come from the optimization of equations expressed in Theorem 1 which are relatively simple; in this case, thanks to the definitions of coefficients b_n and k_n.

Theorem 2. *Suppose that both $U_c(c)$ and $U_p(x)$ are exponential utility functions of the form (5) with $\gamma > 0$. Then, it holds that:*

$$V_{N+1}(x) = 0,$$

$$V_N(x) = \mathbb{P}_N\left(-\frac{1}{\gamma}\right)\exp(-\gamma x) = -k_N\exp\left(-b_N x\right), \tag{12}$$

$$V_N(x) == -k_n\exp\left(-b_n x\right), \quad n \in \{0, 1, \ldots, N-1\},$$

$\forall x \geq 0$, where k_n, b_n, and v_n are given by equations (9)–(11), respectively. In addition, optimal consumption at stage n is $c_n^* = \zeta_n^* x_n$, where

$$\zeta_n^* = \frac{\ln\mathbb{P}_n + b_{n+1}\left(1 + i_{n+1}\right)x_n - \ln\left|k_{n+1}b_{n+1}\left(1 + i_{n+1}\right)v_n\right|}{x_n\left[b_{n+1}\left(1 + i_{n+1}\right) + \gamma\right]}, \tag{13}$$

and the optimal investment at stage n is $a_n^* = S_N^0/S_n^0\left(1 + i_{n+1}\right)\alpha_n^*$, where α_n^* is the solution of (11), which may be found via $\alpha_n^* = 1/\gamma S_n^0/S_N^0\widetilde{\alpha}_n$ where $\widetilde{\alpha}_n$ is the minimum point of

$$\alpha \longmapsto \mathbb{E}\left[\exp\left(-\alpha \cdot R_{n+1}\right)\right], \quad \alpha \in \mathbb{R}^d. \tag{14}$$

Remark 1. Preceding theorem supplies the pursued consumption and investment optimal policy (f_n^*) which is such that

$$f_n^*(x) = \left(\zeta_n^* x_n, a_n^*\right), \tag{15}$$

with ζ_n^* and a_n^* as in Theorem 2.

Proof 1. First of all, consider the following sets of functions in order to test Assumption A.5

$$\mathbb{M} := \{v\colon E \longrightarrow \mathbb{R} \mid v(x) \equiv 0 or v(x) = -k\exp(-bx), \text{for} k, b > 0\},$$
$$\Psi := \left\{f \in F \mid f(x) = (\zeta x, \alpha), \text{with} \alpha \in \mathbb{R}^d, \zeta \in \mathbb{R}_+\right\}. \tag{16}$$

(i) $V_{N+1}(x) \equiv 0$ and $V_N(x) = g_N(x) = U_p(x)$, and both belong to $\mathbb{M}$.

(ii) If $v(x) \equiv 0$, then it is straightforward to see that $\mathcal{T}_n v(x) \in \mathbb{M}$. Assume now that, $v(x) = -k\exp(-bx)$ for $k, b > 0$. In consequence,

$$\begin{aligned}\mathcal{T}_n v(x) &= \sup_{(c,a)\in D(x)}\left\{\mathbb{P}_n U_c(c) + \mathbb{E}v\left[\left(1 + i_{n+1}\right)\left(x - c + a \cdot R_{n+1}\right)\right]\right\}\\ &= \sup_{(c,a)\in D(x)}\left\{\mathbb{P}_n\left(-\frac{1}{\gamma}\right)\exp(-\gamma c) + \mathbb{E}\left[-k\exp\left(-b\left(1 + i_{n+1}\right)\left(x - c + a \cdot R_{n+1}\right)\right)\right]\right\}\\ &= \sup_{(c,a)\in D(x)}\left\{\left(-\frac{\mathbb{P}_n}{\gamma}\right)\exp(-\gamma c) - k\exp\left(-b\left(1 + i_{n+1}\right)(x - c)\right)\mathbb{E}\left[\exp\left(-b\left(1 + i_{n+1}\right)\left(a \cdot R_{n+1}\right)\right)\right]\right\}.\end{aligned} \tag{17}$$

Consider the transformation: $c = \zeta x$ and $a = S_N^0/S_n^0 b(1+i_{n+1})\alpha$, hence:

$$\mathcal{T}_n v(x) = \sup_{0\le\zeta\le1}\left\{\begin{matrix}\left(-\frac{\mathbb{P}_n}{\gamma}\right)\exp(-\gamma\zeta x) - k\exp(-b(1+i_{n+1})(x-\zeta x)) \\ \mathbb{E}_{\alpha\in\mathbb{R}^d}\left[\exp\left(-b(1+i_{n+1})\left(\frac{S_N^0}{S_n^0 b(1+i_{n+1})}\alpha\cdot R_{n+1}\right)\right)\right]\end{matrix}\right\}$$
$$= \sup_{0\le\zeta\le1}\left\{\left(-\frac{\mathbb{P}_n}{\gamma}\right)\exp(-\gamma\zeta x) - k\exp(-b(1+i_{n+1})(x-\zeta x))v_n\right\}, \quad (18)$$

where v_n is given in (11).

(iii) The existence of a maximizer in the set Ψ will be proven. For this, we examine the real function $\ell(\zeta)$ stated as

$$\ell(\zeta) := \left(-\frac{\mathbb{P}_n}{\gamma}\right)\exp(-\gamma\zeta x) - k\exp(-b(1+i_{n+1})(x-\zeta x))v_n. \quad (19)$$

It is possible to discover its maximum through standard optimization techniques. Hence, from Assumption 2, it is observed that the only critical point of ℓ in $[0,1]$ is as follows:

$$\zeta = \frac{\ln \mathbb{P}_n + b(1+i_{n+1})x - \ln\left|kb(1+i_{n+1})v_n\right|}{x\left[b(1+i_{n+1})+\gamma\right]}. \quad (20)$$

which is a relative maximum via the criterion of the second derivative. By substituting the value of ζ in $\mathcal{T}_n v(x)$, it is found that:

$$\mathcal{T}_n v(x) = \left[\left(-\frac{\mathbb{P}_n}{\gamma}\right)\left(\frac{kb(1+i_{n+1})v_n}{\mathbb{P}_n}\right)^{(\gamma/(b(1+i_{n+1})+\gamma))} k\left(\frac{kb(1+i_{n+1})v_n}{\mathbb{P}_n}\right)^{(-b(1+i_{n+1})/(b(1+i_{n+1})+\gamma))} v_n\right]\exp\left(\frac{-b(1+i_{n+1})\gamma}{b(1+i_{n+1})+\gamma}x\right). \quad (21)$$

Therefore, the maximizer for v is of the form $(\zeta x, a)$ and $\mathcal{T}_n v(x) \in \mathbb{M}$

(iv) Expressions for V_n and their corresponding maximizers by utilizing Theorem 1 will be attained.

For each $x > 0$,

$$V_N(x) = \mathbb{P}_N\left(-\frac{1}{\gamma}\right)\exp(-\gamma x) = -k_N\exp(-b_N x). \quad (22)$$

By an inductive process and following essentially the same lines as those in *ii*) and *iiti*), it may be found for $n = 1, \ldots, N-1$ that:

$$V_{N-n}(x) = -k_{N-n}\exp(-b_{N-n}x), \quad \forall x \in E, \quad (23)$$

where k_{N-n} and b_{N-n} are expressed in (9) and (10). Additionally, optimal consumption is given by $c_{N-n} = \zeta_{N-n}x$ with:

$$\zeta_{N-n} = \frac{\ln \mathbb{P}_{N-n} + b_{N-n-1}(1+i_{N-n-1})x - \ln\left|k_{N-n-1}b_{N-n-1}(1+i_{N-n-1})v_{N-n}\right|}{x\left[b_{N-n-1}(1+i_{N-n-1})+\gamma\right]}, \quad (24)$$

and the optimal investment is $a^* = S_N^0/S_{N-n}^0\gamma(1+i_{N-n-1})\alpha^*$, where α^* is the solution of

$$v_{N-n} = \inf_{\alpha^*\in\mathbb{R}^d}\mathbb{E}\left[\exp\left(-\gamma\frac{S_N^0}{S_{N-n}^0}\alpha\cdot R_{N-n-1}\right)\right]. \quad (25)$$

□

Remark 2. The earlier theorem states that under its own assumptions, it is possible to explicitly find the optimal strategy; hence, it is not necessary to perform numerical methods for solving the dynamic programming equation at each stage. However, if this was not the case, there exist several papers dealing with complexity of solution algorithms for MDPs for finite state and action spaces. Nevertheless, we refer to the reader to the contribution of Chow and Tsitsiklis [23] where tight lower bounds on the computational complexity of dynamic programming for the case where the state space is continuous, and the problem is to be solved approximately, within a specified accuracy. On the same direction, Section 12.5 of

[24] is also relevant for the framework studied in this article.

3.2. Numerical Example. In this section, results of Section 3.1 will be illustrated. For this, consider $d = 2$ (two risky assets) and that distribution of relative risk random vectors (R_n) may be approximated by a bivariate normal distribution with parameters μ_n and Σ_n. In this case, it may be found that, for each n

$$v_n = e^{-\mu_{n+1}^T \cdot \Sigma_{n+1}^{-1} \mu_{n+1} + 1/2 \left[\Sigma_{n+1}^{-1}\mu_{n+1}\right]^T \Sigma_{n+1} \left[\Sigma_{n+1}^{-1}\mu_{n+1}\right]}, \tag{26}$$

and that

$$a_n^* = \frac{1}{\gamma(1+i_{n+1})}\Sigma_{n+1}^{-1}\mu_{n+1}. \tag{27}$$

For the sake of simplicity, we consider that random vectors (R_n) are independent and identically distributed with parameters $\Sigma_{n+1} = \begin{pmatrix} 1/8 & 0 \\ 0 & 1/8 \end{pmatrix}$ and $\mu_{n+1} = \begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix}$. Additionally, set $\gamma = 0.6$ for having a not so flat utility function; over the horizontal axis (bigger values of γ lead to utility functions closer to the x-axis), in addition a constant interest rate will be contemplated: $i_n = 0.05$. Finally, it will be established that a random horizon will behave as a binomially distribution with parameters 10 and 0.5.

Now, we split the simulation example into two stages:

3.2.1. Stage I: Before Implementing Dynamics of the Wealth Process. At this stage, it is possible to find corresponding values of v_n and a_n^*. Given the simplifications considered above, it is achievable to find constant values of

$$\begin{aligned} v_n &= 0.877204, \\ a_n^* &= (1.26984, 2.53968). \end{aligned} \tag{28}$$

And the values of b_n and k_n (that will help on construction of optimal consumptions) observed in Table 1.

This is possible since these parameters do not depend on the initial capital or the wealth process.

3.2.2. Stage II: Performing the Wealth Process. At this stage, the initial capital becomes relevant; hence, Tables 2–4 expose the evolution of relative and absolute optimal consumptions as well as a trajectory of (X_n) with an initial capital equal , and . A decreasing behavior is observed on the wealth process accompanied by an increasing behavior of relative consumption, which leads to an "almost" constant performance of absolute consumption. On the same fashion, Figures 1–3 illustrate same observations but allow to compare trajectories of ζ_n, c_n, and x_n.

In Figure 4, the dynamics of relative consumption are observed by considering different values of γ with remaining parameters fixed as before. It may be noticed that practically, no effect is observed.

In Figure 5, sensibility analysis of optimal relative consumption with respect to interest rate is schematized. It is

Table 1: Values of b_n and k_n

n	b_n	k_n
0	0.0740	0.456
1	0.0804	0.5433
2	0.0884	0.5358
3	0.0988	0.4454
4	0.1126	0.3132
5	0.1320	0.1852
6	0.1611	0.0904
7	0.2098	0.0350
8	0.3073	0.0099
9	0.6000	0.0016

Table 2: Evolution of optimal consumptions and wealth processes with $x_0 = \$100$.

n	ζ_n	c	x
0	0.0983	9.83	100
1	0.1255	11.7312	93.4758
2	0.1511	13.2352	87.5927
3	0.1775	13.96	78.6576
4	0.2027	13.6610	67.3956
5	0.2397	14.0671	58.6864
6	0.2816	13.1728	46.7786
7	0.3303	12.5133	37.8847
8	0.4014	10.8936	27.1391
9	0.5199	7.2360	13.8624

Table 3: Evolution of optimal consumptions and wealth processes with $x_0 = \$500$.

n	ζ_n	c	x
0	0.1114	55.7	500
1	0.1238	57.6066	465.3203
2	0.1375	59.1076	429.8738
3	0.1534	59.8102	389.8972
4	0.1729	59.8306	346.0417
5	0.1977	59.9246	303.1092
6	0.2313	59.0432	255.2670
7	0.2798	58.3780	208.6420
8	0.3586	56.7587	158.2786
9	0.5132	53.0666	103.4034

Table 4: Evolution of optimal consumptions and wealth processes with $x_0 = \$1000$.

n	ζ_n	c	x
0	0.1144	114.4	1000
1	0.1233	114.6273	929.6625
2	0.1341	114.7768	855.9050
3	0.1476	114.8599	778.1837
4	0.1650	114.9040	696.3884
5	0.1881	114.8943	610.8150
6	0.2205	114.8113	520.6865
7	0.2691	114.7476	426.4125
8	0.3501	114.5821	327.2842
9	0.5122	114.2206	223.0001

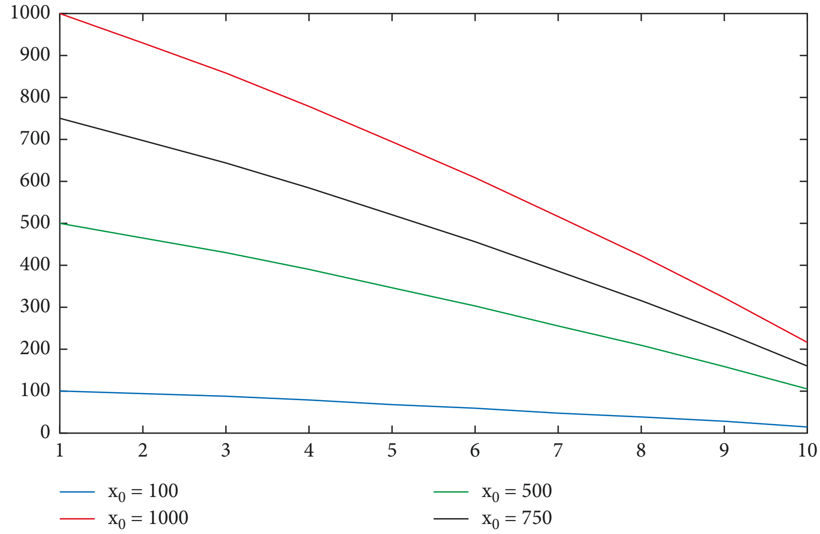

Figure 1: Capital evolution.

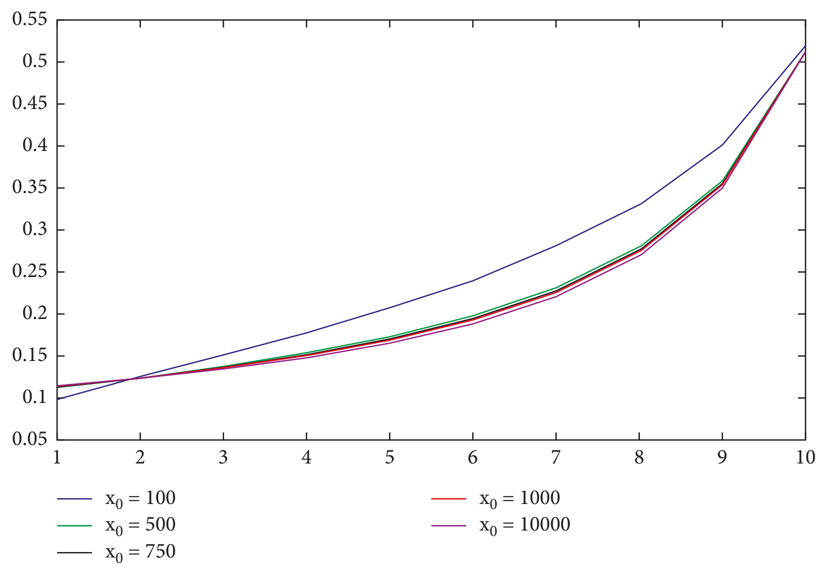

Figure 2: Relative consumption evolution.

possible to appreciate an increasing performance as the constant interest rate increases; nevertheless, Figure 6 exhibits a similar shape on absolute optimal consumptions than Figure 3. In addition, (27) implies a decreasing behavior of optimal investments.

Furthermore, Figure 7 shows a concave behavior on wealth dynamics as interest rate increases, however, one may remark that: naturally, wealth is bigger for huger values of i.

In order to perform an analysis of execution time by implementing the strategy dictated in Theorem 2, an initial capital of with a fixed interest rate of .05 and a horizon with discrete uniform distribution will be taken into account. Hence, in Table 5, the support of random horizon is increased as well as the value of parameter γ. From such a table, together with Figure 8, it is observed that the value of γ has no influence on execution time and that it poses a slow increasing rate (below the identity function) as maximum value of random horizon grows.

On the same fashion, a discretely uniform horizon will be considered again, an initial capital of but now $\gamma = 0.3$ and interest rate will vary from 0.01 to 0.75. In this setting, it is possible to observe in Table 6 as well as in Figure 9 that the interest rate has no repercussion on execution time and that it follows a moderate rise as a random horizon owns a larger support.

By summarizing, it may be seen that neither the value of parameter of exponential utility function nor the interest rate have an effect on performance time and that implementation time grows slower than the maximum value of horizon.

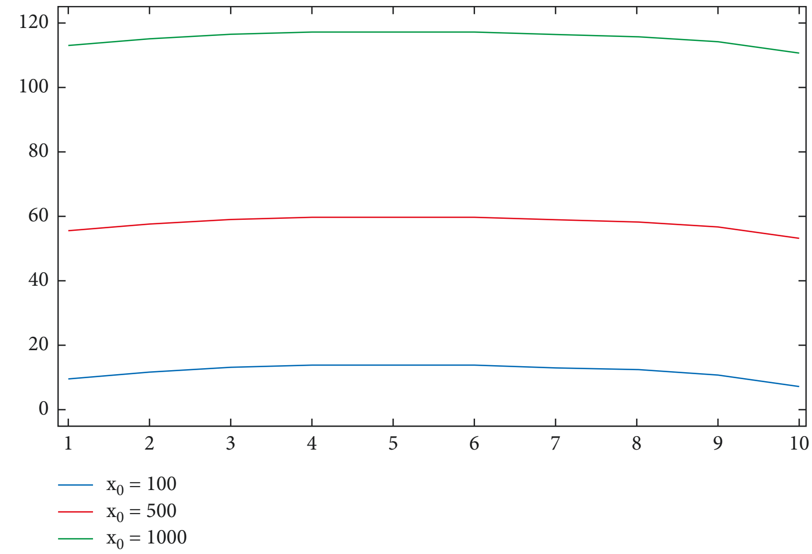

Figure 3: Absolute consumption evolution.

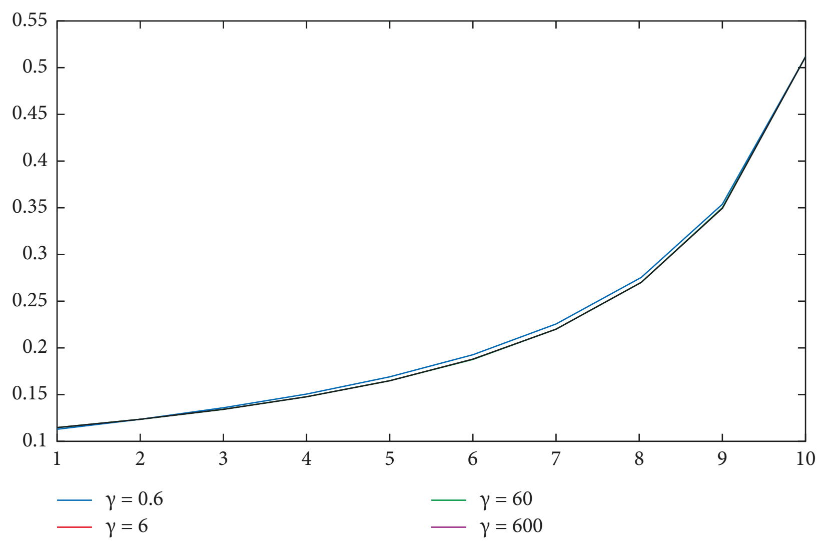

Figure 4: Relative consumption evolution by varying γ.

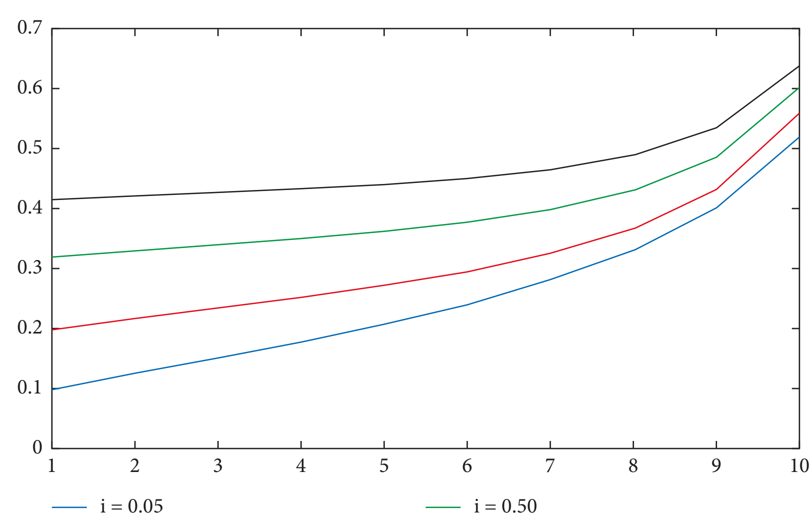

Figure 5: Relative consumption evolution by varying interest rate.

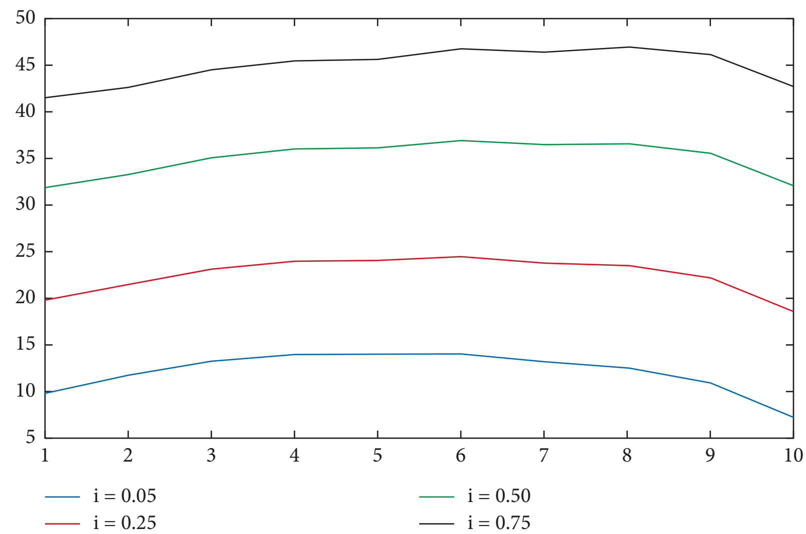

Figure 6: Wealth evolution by varying interest rate.

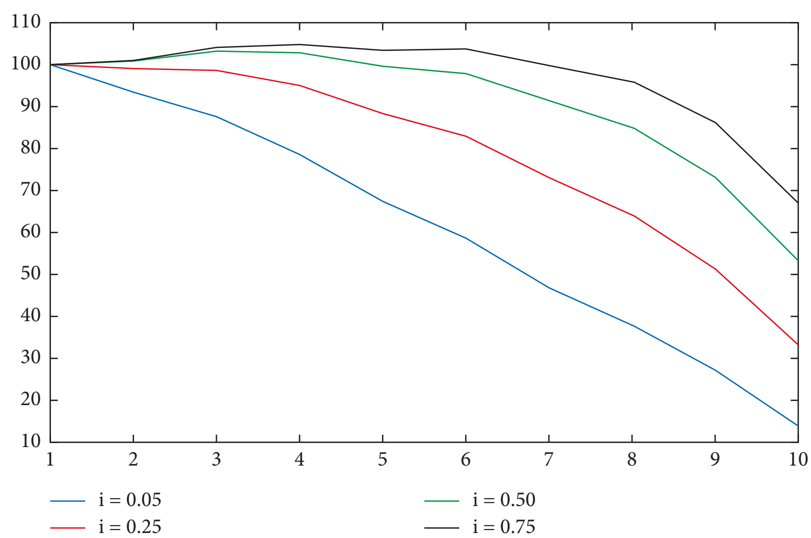

Figure 7: Wealth evolution by varying interest rate.

Table 5: Execution time when the initial capital is $\$1 \times 10^6$ with interest rate $i = 0.05$.

N	$\gamma = 0.1$	$\gamma = 0.5$	$\gamma = 3$	$\gamma = 6$	$\gamma = 60$	$\gamma = 600$
10	0.050062	0.048707	0.054327	0.051612	0.050601	0.053460
100	0.056088	0.054150	0.056277	0.055459	0.058522	0.055845
500	0.099138	0.076228	0.078427	0.076178	0.086510	0.076396
1000	0.100892	0.100931	0.102464	0.123535	0.101802	0.099619
5000	0.310558	0.309741	0.310388	0.310815	0.341130	0.333053
1×10^4	0.618441	0.615198	0.608787	0.637477	0.621824	0.642265
1×10^5	13.093003	13.232268	13.166819	13.687877	13.141081	13.313620
2×10^5	122.116643	119.764554	120.755948	118.576175	117.964425	121.325280
4×10^5	672.962272	672.426565	666.151385	672.633024	691.408851	664.868904
6×10^5	1748.292227	1664.120637	1661.067992	1674.360970	1736.790331	1655.110682
8×10^5	3088.896357	3096.60148	3076.610124	3220.860891	3111.967192	3083.480288
1×10^6	4937.231332	4897.388811	4900.308728	5013.495808	4955.127375	4916.801511

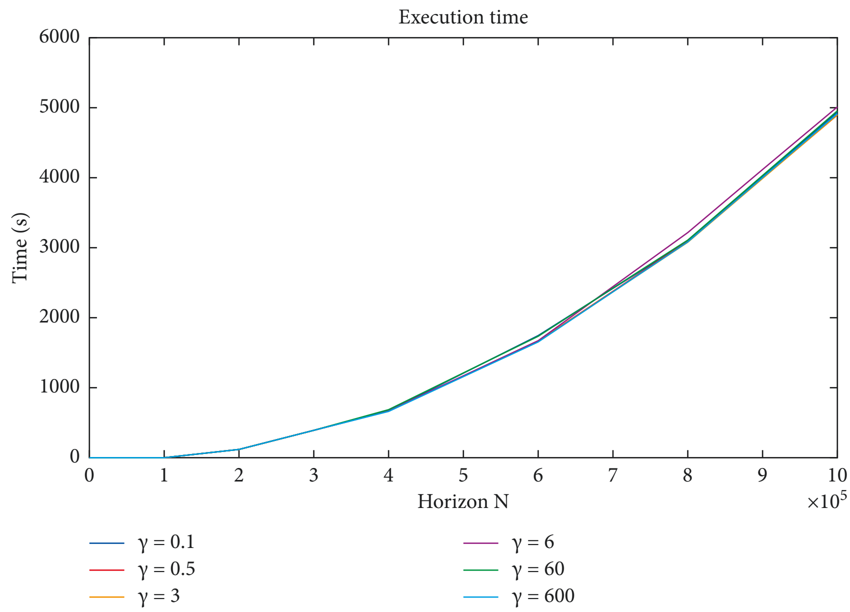

FIGURE 8: Execution time when the initial capital is $1 × 10^6$ with interest rate $i = 0.05$.

TABLE 6: Execution time when the initial capital is $1 × 10^6$ with $\gamma = 0.3$.

N	$i = 0.01$	$i = 0.03$	$i = 0.05$	$i = 0.25$	$i = 0.5$	$i = 0.75$
10	0.049272	0.052132	0.049465	0.049104	0.043670	0.049745
100	0.054473	0.055318	0.057688	0.057867	0.058372	0.055305
500	0.075573	0.077849	0.077799	0.076846	0.076383	0.075571
1000	0.103749	0.101913	0.102075	0.100861	0.103945	0.105100
5000	0.318598	0.309353	0.315017	0.322525	0.324265	0.331267
1×10^4	0.633060	0.664403	0.633980	0.633952	0.635405	0.659529
1×10^5	13.339490	13.377604	13.352550	13.475174	13.379248	13.599737
2×10^5	121.145216	122.143420	122.275267	122.301754	122.777018	126.004807
4×10^5	703.916410	673.852490	677.379424	681.878394	719.429588	678.465843
6×10^5	1667.068711	1678.713265	1664.557272	1692.602557	1678.639482	1716.213704
8×10^5	3152.006812	3139.287981	3125.031973	3083.342423	3087.633951	3092.678243
1×10^6	4883.265378	4877.692400	4873.482225	4885.997006	4992.816111	4870.649850

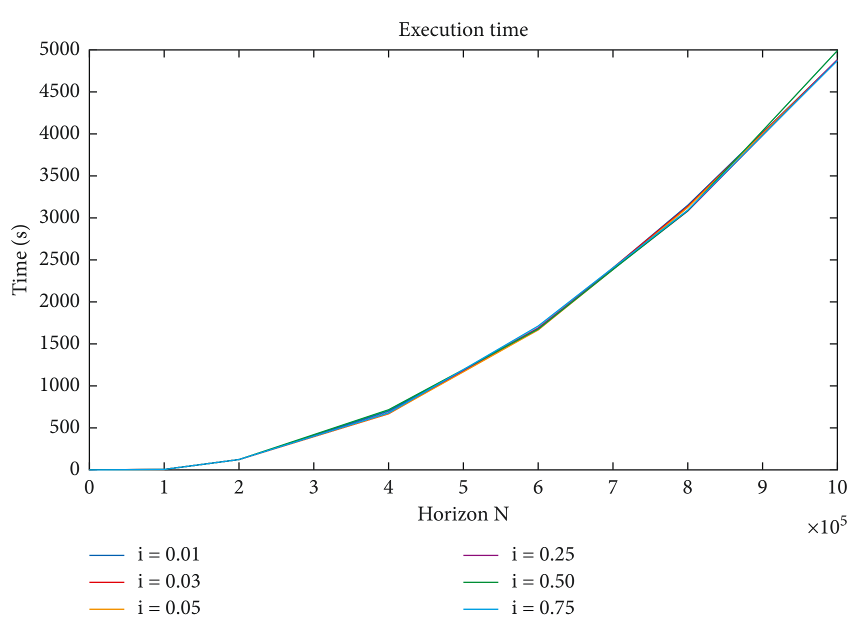

FIGURE 9: Execution time when initial capital is $1 × 10^6$ with $\gamma = 0.3$.

4. Conclusions

In this paper, a consumption and investment problem was studied through a Markov decision process with random horizon of finite support. In this framework, the optimal consumption and investment was obtained via a dynamic programming approach by evaluating consumptions via an exponential utility function.

Appendix

A. Markov Decision Processes with Fixed Horizon

First of all, it will be defined the main topic of this paper: Markov decision processes, which will be the tool to solve the consumption and investment problem described in section 3. As an initial state, horizon will be considered as a fixed natural number. Hence, in such context, the following definition [2] is provided:

Definition A.1. A Markov decision model (MDM) with fixed horizon $N \in \mathbb{N}$ consists of the set (E, A, D, Q, r_n, g) with $n = 0, 1, \ldots, N-1$, where

E is a Borel space, called the state space, endowed with the σ-algebra $\mathfrak{E}$.

A is a Borel space, called the space of actions, equipped with the σ-algebra $\mathfrak{A}$.

$D \subseteq E \times A$ is a measurable subset of $E \times A$ which denotes to the set of possible state-action combinations. It is assumed that D contains the graph of a measurable function $f\colon E \longrightarrow A$. For $x \in E$, the set $D(x) := \{a \in A \mid (x, a) \in D\}$ is called the set of admissible actions.

For each $n \geq 1$, $r_n\colon D \longrightarrow \mathbb{R}$ is a measurable function, which gives the one-stage reward of the system at stage n.

$g\colon E \longrightarrow \mathbb{R}$ is a measurable function, such that $g(x)$ provides the terminal reward if final state is x.

Q is a stochastic transition kernel from D to E. Quantity $Q(B \mid x, a)$ gives the probability that the next state is in B if current state is x and action a is taken.

Remark A.2. Usually, the definition of a Markov decision model considers all its components as time invariant; however, in view of the purposes of this paper, the reward function will depend on time, in fact this condition arises naturally when a random horizon is considered. Additionally, it is also possible to define MDM in a more general setting by considering a time dependence on the set of state-action set, transition functions, and transitions kernels; however, such a paradigm is beyond the interest of this paper, nevertheless, corresponding ideas may be found for example in [14, 25].

In Section 3, transition kernel Q is characterized by random variables (Z_n) defined on some measurable space $(\mathcal{Z}, \mathfrak{Z})$ called the disturbances space. It is assumed that such random variables have a common distribution Q^Z which may depend on $(x, a) \in D$ and that there exists a measurable function $T\colon D \times \mathcal{Z} \longrightarrow E$ known as the transition function. Here, $T(x, a, z)$ provides the next state of the system when the current state is x, action $a \in D(x)$ is taken and disturbance z occurs as follows. Hence, the corresponding transition kernel is defined as follows:

$$Q(B \mid x, a) := Q^Z(\{z \in \mathcal{Z} \mid T(x, a, z) \in B\} \mid x, a), \quad B \in \mathfrak{E}. \tag{A.1}$$

In the context of MDM, decisions are modeled via measurable functions from E to A as can be observed in the following definition.

Definition A.3.

(a) A measurable function $f\colon E \longrightarrow A$, such that $f(x) \in D(x)$ for any $x \in E$, is called the decision rule. Let F denote to the set of all decision rules.

(b) A sequence of decision rules $\pi = (f_0, f_1, \ldots, f_{N-1})$ with $f_n \in F$ is called policy or strategy. The set of this class of policies is denoted by Δ.

One can find more general approaches dealing with policies, a very important reference in that direction is [25]; however, the last definition was adjusted to the intentions of Section 3. The formalization of Markov decision models under a probability space will allow to associate them with some probability measure, and consequently, it will be possible to define the corresponding mathematical expectation.

For this, we contemplate a Markov decision model in N stages, an initial state $x \in E$, $\pi = (f_0, f_1, \ldots, f_{N-1})$ a fixed policy, and the canonical probability space guaranteed by the Ionescu-Tulcea Theorem [14, 25], usually denoted by $(\Omega, \mathcal{F}, \mathbb{P}_x^\pi)$, where $\Omega = E^{N+1}$ and $\mathcal{F}$ is the corresponding product σ-algebra. In addition, if $\omega = (x_0, x_1, \ldots, x_N) \in \Omega$, the state of the system at time n is modeled via a random variable X_n; for $n = 0, 1, \ldots, N$ by

$$X_n(\omega) = X_n(x_0, x_1, \ldots, x_N) = x_n. \tag{A.2}$$

On this probability space, (X_n) is called the Markov decision process. Given that, the optimization problems that are treated in this article are related with optimization of expected values of aggregated rewards, the following assumption [14] is considered:

Assumption A.4. For $x \in E$, $\sup_{\pi \in \Delta} \mathbb{E}_x^\pi [\sum_{k=0}^{N-1} r_k^+ (X_k, f_k(X_k)) + g^+(X_N)] < \infty$.

In all the sequel, it will supposed that Assumption A.4 holds for a Markov decision process with horizon N.

Contemplated performance criteria of policy π when initial state is $x \in E$ is the so-called total expected reward:

$$V(\pi, x) := \mathbb{E}_x^\pi \left[\sum_{k=0}^{N-1} r_k(X_k, f_k(X_k)) + g^+(X_N) \right], \quad x \in E. \tag{A.3}$$

Then, the value function for $x \in E$ is defined by the following:

$$V^*(x) := \sup_{\pi\in\Delta} V(\pi, x), \quad x \in E. \tag{A.4}$$

Functions $V(\pi, x)$ and $V(x)$ are well defined since

$$V(\pi, x) \le V^*(x) \le \sup_\pi \mathbb{E}_x^\pi \left[\sum_{k=0}^{N-1} r_k(X_k, f_k(X_k)) + g^+(X_N)\right] < \infty. \tag{A.5}$$

A policy π^* is called optimal for a N stage MDP, if $V(\pi^*, x) = V^*(x)$ for all $x \in E$, [2, 25].

The following assumption allows to provide sufficient conditions to establish the existence of optimal policies [14].

Assumption A.5. There exist sets $\mathbb{M} \subset \mathbb{M}(E) := \{v: E \longrightarrow [-\infty, \infty) | v \text{ is measurable}\}$. and $\Psi \subset F$, for $n = 0, 1, \ldots, N-1$, such that:

(i) $g \in \mathbb{M}$.

(ii) If $v \in \mathbb{M}$, then $\mathcal{T}_n v(x)$ is well defined and $\mathcal{T}_n v \in \mathbb{M}$. (Here $\mathcal{T}_n v(x) := \sup_{a\in D(x)}\{r_n(x, a) + \int v(x') Q(dx'|x, a)\}$, $x \in E$).

(iii) For all $v \in \mathbb{M}$ there exists a maximizer $f \in \Psi$ of v; i.e.$\mathcal{T}_n v(x) = r_n(x, f(x)) + \int v(x')Q(dx'|x, f(x))$, $x \in E$.

The dynamic programming technique is expressed in the following theorem, whose proof may be found; for example, in [2, 14, 25].

Theorem A.6. *Let $V_0, V_1, \ldots, V_N$ be functions on E defined by the following:*

$$V_N(x) := g(x), \tag{A.6}$$

and for $n = N-1, N-2, \ldots, 0$

$$V_n(x) := \sup_{a\in D(x)} \left[r_n(x, a) + \int_E V_{n+1}(y) Q(dy|x, a)\right]. \tag{A.7}$$

Suppose that Assumption A.5 holds, then there exist maximizers $f_n \in \Psi$ of V_n, in addition the deterministic Markov policy $\pi^* = (f_0, \ldots, f_{N-1})$ is optimal, and the value function V^* equals V_0, i.e.,

$$V^*(x) = V_0(x) = V(\pi^*, x) \forall x \in E. \tag{A.8}$$

B. Financial Markets. Financial markets allow an efficient allocation of resources within the economy. Through organized and regulated exchanges, these markets will give to participants a certain guarantee that they will be treated fairly and honestly. In short, it is a platform that allows traders to easily buy and sell financial instruments and securities, for example, stocks, bonds, commercial paper, bills of exchange, debentures, and more. Financial markets lie in the fact that they act as an intermediary between savers and investors, or they help savers to become investors [26, 27].

It will be considered a financial market of N-periods with d risky assets and a risk-free bond with considerations treated in [14]. It will be assumed that random variables are defined in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with a filtration $(\mathcal{F}_n)$ with $\mathcal{F}_0{:} = \{\emptyset, \Omega\}$. The financial market is given by the following:

A risk-free bond with $S_0^0 \equiv 1$ and

$$S_{n+1}^0 := S_n^0 (1 + i_{n+1}), \quad n = 0, 1, \ldots, N-1 \tag{A.9}$$

where i_{n+1} denotes the deterministic interest rate for period $[n, n+1)$.

There are d risky assets and the price process of k-th asset is given by $S_0^k = s_0^k$ known and

$$S_{n+1}^k = S_n^k \widetilde{R}_{n+1}^k, \quad n = 0, \ldots, N-1, \tag{A.10}$$

Where $\widetilde{R}_{n+1}^k > 0$ $\mathbb{P}$-a.s. for all k and n. $\widetilde{R}_{n+1}^k$ is the relative price change on interval $[n, n+1)$ for k-th risky asset and process (S_n^k) is assumed to be adapted to $(\mathcal{F}_n)$ for any k.

Positive random variable $\widetilde{R}_{n+1}$ defines the relative price change S_{n+1}/S_n. Relative risk process (R_n) is defined by $R_n := (R_n^1, \ldots, R_n^d)$ and $R_n^k := \widetilde{R}_n^k / 1 + i_{n+1} - 1$, $k = 1, \ldots, d$.

Consider now the following notation, $S_n := (S_n^1, \ldots, S_n^d)$, $\widetilde{R}_n := (\widetilde{R}_n^1, \ldots, \widetilde{R}_n^d)$ and $\mathcal{F}_n^S := \sigma(S_0, \ldots, S_n)$. As (S_n) is $(\mathcal{F}_n)$ adapted it holds that: $\mathcal{F}_n^S \subset \mathcal{F}_n$ for $n = 0, 1, \ldots, N-1$. It is assumed that $(\mathcal{F}_n)$ is the filtration generated by stock prices, that is $\mathcal{F}_n = \mathcal{F}_n^S$. Subsequent definition is the main needed mathematical object for investing in the earlier described financial market.

Definition B.7. A portfolio or trading portfolio is a stochastic process $(\mathcal{F}_n)$-adapted, $\phi = (\phi_n^0, \phi_n)$ where $\phi_n^0 \in \mathbb{R}$ and $\phi_n = (\phi_n^1, \ldots, \phi_n^d) \in \mathbb{R}^d$ for $n = 0, 1, \ldots, N-1$. Random variable ϕ_n^k denotes the amount of money invested in k-th asset during $[n, n+1)$.

Therefore, the wealth process evolves as follows:

$$X_{n+1} = (1 + i_{n+1})(X_n + \phi_n \cdot R_{n+1}), \tag{A.11}$$

In order to solve the consumption and investment problem of Section 3, an utility function for evaluating consumptions will be needed.

Definition B.8. A function, $U\colon \operatorname{dom} U \longrightarrow \mathbb{R}$ is called a utility function, if U is strictly increasing, strictly concave and continuous on its domain.

The following assumption correspond to the proper version of Assumption A.4 in the Financial market context.

Assumption B.9. $\mathbb{E}\|R_n\| < \infty$. Where $\|R_n\| = |R_n^1| + \ldots + |R_n^d|$, for $R_n \in \mathbb{R}^d$ and $n = 1, \ldots, N$.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete-Time Case*, Athena Scientific, Belmont, MA, USA, 1978.

[2] O. Hernández-Lerma and J. B. Laserre, *Further Topics on Discrete-Time Markov Control Processes*, Springer-Verlag, Berlin, Germany, 1999.

[3] K. Hinderer, *Foundations of Non-sationary Dynamic Programming with Discrete-Time Parameter*, Springer-Verlag, Berlin, Germany, 1970.

[4] G. Arindam and S. Biswajit, *Economically Independent Reverse Logistics of Customer-Centric Closed-Loop Supply Chain for Herbal Medicines and Biofuel*, Elsevier, vol. 334, Amsterdam, Netherland, Article ID 129977, 2022.

[5] G. Arindam, "Fractile criterion iterative-interactive optimisation process for multi-objective stochastic linear programming problems in fuzzy environment," *International Journal of Mathematics in Operational Research*, vol. 18, no. 3, pp. 289–309, 2021.

[6] G. Arindam and K. R. Tapan, "Multi-objective optimization of cost-effective and customer-centric closed-loop supply chain management model in T-environment," *Soft Computing*, vol. 24, 2020.

[7] S. Biswajit, M. Arunava, S. Mitali, K. D. Bikash, and R. Gargi, "Two-echelon supply chain model with manufacturing quality improvement and setup cost reduction," *Journal of Industrial and Management Optimization*, vol. 13, no. 2, pp. 1085–1104, 2017.

[8] K. Pentikousis, O. Blume, R. A. Calvo, and S. Papavassiliou, "Mobile networks and management," *Social-Informatics and Telecommunications Engineering*, Springer, Berlin, Germany, Lecture Notes of the Institute for Computer Sciences, 2009.

[9] H. Cruz-Suárez, R. Ilhuicatzi-Roldán, and R. Montes-de-Oca, "Markov decision processes on Borel spaces with total cost and random horizon," *Journal of Optimization Theory and Applications*, vol. 162, no. 1, pp. 329–346, 2012.

[10] V. Vázquez-Guevara, H. Cruz-Suárez, and F. Velasco-Luna, "Optimal assignment of sellers in a store with a random number of clients via the armed bandit model," *RAIRO-Operation Research*, vol. 51, no. 4, pp. 1119–1132, 2017.

[11] K. J. Chung and M. J. Sobel, "Continuous-time Markov decision processes with exponential utility," *SIAM Journal on Control and Optimization*, vol. 25, no. 1, pp. 49–62, 1987.

[12] K. Yoshinobu, K. Masami, and Y. Masami, "Discounted Markov decision processes with utility constraints," in *Computers and Mathematics with Applications*, vol. 51, pp. 279–284, Elsevier, Amsterdam, Netherland, 2006.

[13] S. Carpin, Y.-L. Chow, and M. Pavone, "Risk aversion in finite Markov Decision Processes using total cost criteria and average value at risk," in *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 335–342, Stockholm, Sweden, May 2016.

[14] N. Bäuerle and U. Rieder, *Markov Decision Processes with Aplications to Finance*, Springer-Verlag, Berlin, Germany, 2001.

[15] T. Bodnar, N. Parolya, and W. Schmid, "On the exact solution of the multi-period portfolio choice problem for an exponential utility under return predictability," *European Journal of Operational Research*, vol. 246, no. 2, pp. 528–542, 2015.

[16] T. Bodnar, D. Ivasiuk, N. Parolya, and W. Schmid, "Mean-variance efficiency of optimal power and logarithmic utility portfolios," *Math Finance Economics*, vol. 14, pp. 675–698, 2020.

[17] Z. Nika and M. Rásony, "Log-optimal portfolios with memory effect," *Applied Mathematical Finance*, vol. 25, no. 5-6, pp. 557–585, 2018.

[18] R. A. Jarrow, *Continuous-Time Asset Pricing Theory*, Springer Finance, New York, NY, USA, 1st edition, 2018.

[19] H. U. Gerber and G. Pafumi, "Utility functions: from risk theory to finance," *North American Actuarial Journal*, vol. 2, no. 3, pp. 74–91, 1998.

[20] N. Bäurle and U. Rieder, "More risk-sensitive Markov decision processes," *Mathematics of Operations Research*, vol. 39, no. 1, pp. 105–120, 2014.

[21] M. Bouakiz and M. J. Sobel, "Inventory control with an exponential utility criterion," *Operations Research*, vol. 40, no. 3, pp. 603–608, 1992.

[22] Y. Zhang, "Continuous-time Markov decision processes with exponential utility," *SIAM Journal on Control and Optimization*, vol. 55, no. 4, pp. 2636–2660, 2017.

[23] C. S. Chow and J. N. Tsitsiklis, "The complexity of dynamic programming," *Journal of Complexity*, vol. 5, pp. 466–488, 1989.

[24] A. L. Almudevar, *Approximate Iterative Algorithms*, CRC Press, Florida, FL, USA, 1st edition, 2014.

[25] O. Hernández-Lerma and J. B. Laserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Springer-Verlag, Berlin, Germany, 1996.

[26] R. L. Heilbroner and W. Milberg, *The Making of Economic Society*, Prentice-Hall, New Jersey, NJ, USA, 13th edition, 2011.

[27] J. Madura, *Financial Markets & Institutions*, Cengage Learning Inc., USA, 13th edition, 2020.