Business and Management Studies
Vol. 9, No. 2; December 2023
ISSN: 2374-5916 E-ISSN: 2374-5924
Published by Redfame Publishing
URL: http://bms.redfame.com

Hedging Interest Rate Options with Reinforcement Learning: an investigation of a heavy-tailed distribution

Allan Jonathan da Silva[1,2], Jack Baczynski[2], Leonardo Fagundes de Mello[2]

[1] Federal Center for Technological Education - CEFET/RJ, Itaguaí – RJ, Brazil

[2] National Laboratory for Scientific Computing - LNCC, Petrópolis, Brazil

Correspondence: Allan Jonathan da Silva, Department of Production Engineering, CEFET/RJ, 23812-101 Itaguaí-RJ, Brazil. E-mail: allan.jonathan@cefet-rj.br

Received: November 8, 2023 Accepted: December 1, 2023 Available online: December 20, 2023

doi:10.11114/bms.v9i2.6515 URL: https://doi.org/10.11114/bms.v9i2.6515

Abstract

Purpose: The study intends to model an interest rate index option using a heavy-tailed distribution. The goal is to calculate the interest rate path-dependent option prices that are consistent with market data and to develop a reinforcement learning strategy to discretely hedge the position considering transaction costs. Methodology: This paper presents a mathematical framework to calculate the price of interest rate path-dependent options. The research adapted a Fourier cosine series formula to employ the characteristic function of the present value of the forward index, which is modeled as a variance-gamma process and uses deep Q-learning to hedge such options. Findings: There is market evidence that the implied volatility curve is not flat. The study demonstrated that the variance-gamma process generates an increasing volatility smile, which is consistent with market observations. Additionally, hedging results show that the path-dependent options generated from the variance-gamma process can be efficiently hedged with advanced Q-learning techniques. Research limitations/implications: The study comprised only the variance-gamma process. Other probability distributions, such as the Normal Inverse Gaussian model, should be investigated. Practical implications: This study reveals which type of probability distribution should be present in a pricing engine to be consistent with implied volatilities. The approach provided here can assist managers in evaluating and comprehending market pricing behavior as well as achieving discrete hedging with costs. Originality: The paper addressed the merging of a fast pricing method for the interest rate options with a heavy-tailed distribution and the discrete interest rate derivatives hedging with reinforcement learning.

Keywords: interest rate options, COS method, reinforcement learning, heavy-tail probability distributions, discrete hedging.

1. Introduction

The derivatives market plays an important role in the global financial landscape, enabling a wide range of financial operations and risk management measures. Derivatives are financial products that derive their values from underlying assets or variables such as commodities, currencies, interest rates, or stock indexes (see Smithson (1998), Fabozzi (2000) and Martellini et. al (2003)). This class of financial contract allows participants to hedge against price fluctuations, speculate on market movements and gain exposure to various asset classes without directly owning the underlying assets (Hull (2009)). Derivatives are characterized by their flexibility, diversity and potential for both profit and risk (Marshall and Bansal (1991)). Derivatives contracts offer opportunities for investors to optimize their portfolios, manage uncertainties and tailor financial strategies to their specific needs.

According to Marshall and Bansal (1991), financial derivative contracts can be divided into the following categories:

- Futures: Agreement to buy or sell an underlying asset at a future date at an agreed-upon price. Used for speculation or hedging price risks. Standardized and traded on exchanges.
- Options: Gives the buyer the right, but not the obligation, to buy (call) or sell (put) an asset at a predetermined price. Premium is paid by the buyer to the option seller. Used for speculation, hedging, or income strategies.
- Swaps: Exchange of cash flows between parties based on underlying assets like interest rates or currencies. Used to manage interest rate, foreign exchange, or other risks.

- Forwards: Similar to futures but traded over-the-counter (OTC). Can be customized to particular needs. Credit risks are more significant compared to futures.
- Exotic Options Contracts: Variations of standard options with customized features. Examples include barrier options, lookback options, etc. Used in specific situations to meet complex hedging or investment needs.

The above list provides insight into the multifaceted derivatives market, highlighting its role as a catalyst for financial innovation, risk mitigation and portfolio diversification. As we continue to explore this topic, this paper will delve deeper into an exotic type of derivative contract: a Asian interest rate option commonly traded in the Brazilian interest rate market (see Wilmott (2006) for a mathematical definition of Asian options).

The existing singularities in the Brazilian market stirred the development of particular derivatives contracts. Among them, we have a singular kind of interest rate option, namely the Interbank Deposit rate Index (IDI) option. The IDI option is the main interest rate option offered by B3 (the Brazilian stock, futures and derivatives exchange). It is of European type with cash settled at maturity (see Hull (2009) for types of option exercising). The underlying variable of the option, the IDI, is an index that accumulates from an initial value according to the daily Interbank Deposit rate (DI rate). The IDI option is an Asian option since its payoff depends on the path of the DI rate (or the interest rate) during the life of the option. Among other advantages, Asian options reduce the risk of market manipulation of the underlying instrument at maturity. Moreover, according to Hull (2009), the average feature makes Asian options typically cheaper than vanilla options. Detailed discussions about the IDI options can be found in da Silva et. al (2016) and Carreira and Brostowicz (2016).

Similar path-dependent products, in a mathematical sense, are commonly found in commodities and currency markets, as an average rate contract. An example of the usage of IDI option is an institutional investor or a bank with a floating rate debt seeking to protect itself against an unexpected rise in interest rates.

In the domain of the mathematical modeling, the Gaussian distribution, also known as the normal distribution, is essential in option pricing models and quantitative finance (see Neftci (2000)). This symmetrical bell-shaped curve is a fundamental assumption in various pricing models, including the popular Black-Scholes model (Black and Scholes (1973)). Its importance lies in its ability to simplify the complex nature of financial markets by capturing the behavior of asset returns under certain assumptions, as the one that asset returns follow a log-normal distribution, which results in normally distributed changes in asset prices over time. The Black-Scholes model is adapted to the futures market through the Black-76 model (Black (1976)). According to Carreira and Brostowicz (2016), the Black-76 model is the standard model used to price IDI option and quote implied volatilities.

In the interest rate market, the Vasicek model assumes that the terminal distribution of interest rates is normally distributed (Vasicek (1977)). As can be seen in Brigo and Mercurio (2006) and Fabozzi (2000), under the normality assumption, the Vasicek model generates closed formula for the price of a wide range of fixed income derivatives. This includes the IDI option pricing (see Vieira and Pereira (2000)). As can be seen, the normal distribution serves as a cornerstone in option pricing models by simplifying the complex behavior of asset returns.

However, heavy-tailed distributions, a fascinating phenomenon within the realm of probability and statistics, have garnered significant attention in the field of finance due to their profound implications for risk assessment, asset pricing, and portfolio management (see e.g. Vellekoop and Nieuwenhuis (2007), De Domenico et. al (2023) and Jondeau et. al (2007)). Heavy-tailed distributions display a deviation from the conventional Gaussian distribution as they exhibit a higher frequency of extreme events compared to what traditional models would predict. This atypical trait in heavy-tailed distributions challenges conventional assumptions and underscores the need for a deeper understanding of risk in financial markets. A gentle introduction to the financial modeling of non-Gaussian distributions can be found in Zhu (2009). A more technical treatment can be found in Tankov and Cont (2003).

In Madan et. al (1998) is proposed the variance-gamma process, a three parameter stochastic process that generalizes Brownian motion. It was applied for modeling the dynamics of log stock prices by assessing Brownian motion with drift at a random time determined by a gamma process. The authors established a closed form formula for the vanilla option price and demonstrated that the risk neutral density for a database of historical prices is negatively skewed with a larger kurtosis.

In finance, the presence of heavy tails has far-reaching consequences that extend beyond theoretical considerations:

- Tail Risks: Heavy-tailed distributions emphasize the importance of tail risks - extremely rare but impactful events that can lead to substantial market disruptions (see e.g. Tankov and Cont (2003)).
- Implied Volatility: Heavy-tailed distributions contribute to observed patterns of market volatility. The increased likelihood of extreme market moves is reflected in options pricing, volatility smiles, and skewness in market data (see Gatheral (2006) and Brigo and Mercurio (2006)).

- Risk Management: Understanding and quantifying the risks associated with heavy-tailed distributions is crucial for effective risk management. The potential for rare but severe events can impact the stability and performance of investment portfolios (see Khalaf et. al (2021)).
- Derivatives pricing: The recognition of heavy tails prompts a reevaluation of these models and the development of more robust frameworks that accurately capture the tail behavior of asset returns (see e.g. Bouziane (2008)).

Heavy-tailed distributions also hold significant importance in the IDI option pricing due to their ability to provide valuable insights into extreme market events - a suddenly hike in the interbank interest rate target - that can significantly impact portfolios and investment strategies. These distributions exhibit a tendency to generate larger-than-anticipated observations in their tails, thereby challenging the traditional assumptions in various mathematical models commonly employed in the financial domain. According to Ornelas and Takami (2011), the risk-neutral densities implied on IDI options exhibited that, the Brazilian interest rates also presents non-Gaussian distribution. Under the Vasicek model, an IDI option price formula can be found in Vieira and Pereira (2000). In Barbachan and Ornelas (2003), the IDI option is calculated under the CIR model (Cox et al. (1986)). Junior et. al (2003) and Almeida et. al (2003) assumed that the short-term rate follows the Hull and White (1993) model and obtained analytical solutions for the price of IDI options. The HJM model is implemented in Barbedo et. al (2010) to price IDI options. The problem was also numerically solved via a finite difference method in da Silva et. al (2016). Genaro and Avellaneda (2018) showed that the price of the IDI option is sensitive to changes in monetary policy and da Silva et. al (2019) enhanced the Vasicek model with exponential, gamma and normal jumps.

Although the vast existing literature cited theretofore and others not mentioned, to the best of our knowledge there is no analytical or numerical model used to calculate the prices of IDI Options with Lévy process (Tankov and Cont (2003)). In order to capture the heavy-tail behavior observed in the IDI option market prices, this study aims to address the gap of computational methods enhancing the standard Black-76 model to price IDI options. The numerical method employed in this study has emerged recently in the literature. The adaptation of the COS method (Fang and Oosterlee (2008)) to interest rate models can be seen in the works da Silva et. al (2019), da Silva et. al (2020) and da Silva et. al (2023).

We also face the problem of discretely hedging financial options. While the Black-Scholes and the Black-76 models provide a robust framework for option pricing in idealized conditions, it does not account for real-world constraints and frictions (Hull (2009)). In practice, financial markets are characterized by discrete trading, transaction costs, and limitations on the frequency of hedging adjustments. This leads to a significant gap between the idealized continuous hedging strategy implied by the Black-Scholes formula and the practical challenges faced by traders and investors (Smithson (1998)). As such, a crucial hypothesis underlying this model is its smooth, continuous, and costless hedging approach, which does not align with the complexities of real markets. Addressing these issues is essential for improving the accuracy and effectiveness of option pricing and hedging strategies in the dynamic and imperfect world of financial markets.

In this work, we will harness the power of artificial intelligence in the domain of financial instrument hedging, with a particular focus on fixed-income derivatives. Our objective is to explore how AI-driven techniques can enhance hedging strategies in the realm of fixed-income financial instruments. By leveraging machine learning models, we aim to optimize risk mitigation and portfolio performance, ultimately contributing to more effective and precise hedging practices in the complex world of fixed-income derivatives.

Artificial intelligence (AI) is undeniably reshaping the landscape of finance and investment management. From leveraging genetic algorithms for investment strategy optimization (Núñez-Letamendia (2002)) to harnessing neural networks for customer deposit predictions (Gafrej (2023)), AI is not only improving predictive accuracy but also introducing new dimensions of efficiency, convenience, and security. It is fostering innovation in pricing models, risk management, and even redefining traditional roles in wealth management (Nain and Rajan (2023)). The findings of Dandapani (2017) consistently emphasize the potential for AI to revolutionize the field, encouraging a more sophisticated, data-driven approach, and fostering a deeper understanding of financial markets. As AI technology continues to evolve, it is poised to play an increasingly vital role in the future of finance and investment management.

This paper is organized as follows: in the Section 2 we present the mathematics of IDI options, its payoff function and the standard pricing formula used by market participants to quote the implied volatility. We also present the Fourier cosine series method (COS), which is used to calculate the probability density function numerically and the reinforcement learning subject. In Section 3, we present the variance-gamma process and show how the COS method is adapted to the IDI option pricing. We also show how the Q-learning method is used to develop a hedging strategy for the IDI options. In Section 4, we show the common shape of the implied probability distribution from IDI option prices.

Later, we compare the prices given by the Black-75 formula and that calculated via the COS method. We also present the variance-gamma process, the statistical model proposed in this paper to capture the heavy tail observed in the IDI option prices. Finally, we present the dynamic hedging performed with reinforcement learning. Section 5 concludes the paper.

2. Literature Review

2.1 The mathematics of IDI options

We assume an interest rate market with underlying probability space $(\Omega, \mathbb{F}, \mathbb{Q})$ equipped with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ where $\mathbb{Q}$ is the risk neutral measure. According to the B3 protocols, the DI rate is the average of the interbank rate of a one-day-period, calculated daily and expressed as the effective rate per annum.

So, the ID index (IDI) accumulates discretely, according to

$$y(T) = y(t)\prod_{j=1}^{t-1} \left(1 + DI_j\right)^{\frac{1}{252}} \tag{1}$$

where j denotes the end of day and DI_j assigns the corresponding DI rate.

If we approximate the continuously DI rate by the instantaneous continuously compounding interest rate, i.e. $r(t) = \ln\left(1 + DI(t)\right)$, the index can be represented by the following continuous compounding expression

$$y(T) = y(t)e^{\int_t^T r(s)ds}. \tag{2}$$

Given non-negative interest rates, the index is a non-decreasing function of r(s).

A European call option is a contract that gives the owner the right, but not the obligation, to buy a specified amount of an underlying security at a specified price and at a specified time. The payoff of the IDI call option maturing at T is

$$\max(y(T) - K, 0), \tag{3}$$

where K is the strike price. Therefore, the price at time t of this option is

$$C(t,T) = \mathbb{E}\left[e^{-\int_t^T r(s)ds}\max(y(T) - K, 0) \mid \mathcal{F}_t\right]. \tag{4}$$

The forward ID index is given by

$$F(t) = \mathbb{E}\left[y(t)e^{\int_t^T r(s)ds} \mid r(t)\right]. \tag{5}$$

Note that F(T)=y(T). Then the price at time t of the IDI option with strike K and maturity T is

$$\begin{aligned} C(t,T) &= \mathbb{E}\left[e^{-\int_t^T r(s)ds}\max(F(T) - K, 0) \mid \bar{y}(t)\right] \\ &= \mathbb{E}\left[\max\left(e^{-\int_t^T r(s)ds}F(T) - \bar{K}, 0\right) \mid \bar{y}(t)\right] \\ &= \mathbb{E}[\max(\bar{y}(T) - \bar{K}, 0) \mid \bar{y}(t)], \end{aligned} \tag{6}$$

where the term $\bar{y}(T) = e^{-\int_t^T r(s)ds}F(T)$ is called the present value of the forward ID index and $\bar{K} = Ke^{-\int_t^T r(s)ds}$ is the discounted strike price. Conversely, European put options give holders of the option the right, but not the obligation, to sell a specified amount of an underlying security at a specified price and at a specified time. The payoff of the IDI put option maturing in T is

$$\max(K - y(T), 0). \tag{7}$$

By put-call parity, the price at time t of the IDI put option is

$$G(t,T) = C(t,T) + KD(t,T) - y(t), \tag{8}$$

where D(t,T) is the zero-coupon bond price with maturity in T.

2.2 Analytical results

In this section we present some analytical results found in the literature for IDI option prices built upon classical models.

For some probability distributions, it is possible to solve the risk neutral expectation (4) without resorting to numerical methods.

Closed-form formulas are popular among practitioners despite their distance to realistic assumptions. The main attraction for analytical solutions is that the outcomes of the model are widely known and the formula can be easily implemented in a simple spreadsheet.

Let us assume that the present value of the forward IDI follows the geometric Brownian motion as below

$$d\bar{y}(t) = \sigma \bar{y}(t) dW(t), \tag{9}$$

where σ is the volatility and W(t) is the standard Wiener processes. The above process is known as the Black-76 model, due to Black (1976). The call price formula is given by

$$\begin{aligned} C(t,T) &= D(t,T)[y(t)N(d_1) - KN(d_2)], \\ d_1 &= \frac{\ln \frac{y(t)}{K} + \frac{\sigma_y^2 T}{2}}{\sigma_y \sqrt{T}}, \\ d_2 &= \frac{\ln \frac{y(t)}{K} - \frac{\sigma_y^2 T}{2}}{\sigma_y \sqrt{T}} \end{aligned} \tag{10}$$

where D(t,T) is the discount factor, y(t) is the spot index and N(·) denotes the cumulative standard normal distribution function. The resulting probability density function of the forward index is the lognormal distribution. The model shown above assumes that the index is independent of the interest rate, which is considered to be constant, and that volatility is similarly constant over time. Despite these implausible assumptions, the market chose the Black-76 model to estimate option pricing due to its simplicity. Carreira and Brostowicz (2016) contains technical details and debate.

Additional analytical models for the IDI call option problem can be found in the literature. Under the Vasicek model (Vasicek (1977)) hypothesis for the short rate, Vieira and Pereira (2000) developed a closed-form solution for this pricing problem. Barbachan and Ornelas (2003) assumed that the interest rate follows the one-factor square-root process, known as CIR due to Cox et al. (1986). Junior et. al (2003) and Almeida et. al (2003) assumed that the short-term rate follows the Hull and White (1993) process, which is an one-factor process too. In order to model the instantaneous forward rates, Barbedo et. al (2010) implemented a three-factor HJM model (Heath et. al (1992)) to price IDI options. Almeida and Vicente (2012) implemented a Gaussian model to study the movements of the term structure of the interest rates. Finally, Genaro and Avellaneda (2018) extend a Gaussian model with jumps at deterministic times.

Realistic models often suffer from the lack of analytical tractability. In order to improve the model capabilities with random jumps, stochastic volatilities and correlation between stochastic variables, we need to resort to numerical methods. Among the numerical methods, Finite-difference, Monte Carlo simulation and Fourier-transform are the most common techniques used to find the price of a contingent claim.

The Finite-difference method aims to solve the partial differential equation resulted from the application of the Feynman-Kac theorem in the payoff formula. da Silva et. al (2016) solves the IDI option pricing problem via a modified full implicit finite difference method using the Vasicek model.

Monte Carlo methods aims to simulate thousands of paths of the underlying source of risk in order to estimate the mean value of the discounted payoff. A nice reference for this approach is Glasserman (2004). It is well known that this technique suffers of the curse of dimensionality.

Fourier-transform based approaches seeks to find - at least numerically, the characteristic function of the random variable that underlies the financial derivative. Applications of this approach in interest rate markets can be found in Duffie (2008), Zhu (2009) and Bouziane (2008). In the next sections we will apply the Fourier-cosine series approach presented in Fang and Oosterlee (2008) in order to find the probability density function under concern.

2.3 Recovering probability functions with Fourier series: the COS method

In some cases, the probability density or mass function is not available in closed form. Classical distributions (e.g., normal and lognormal densities) that lead to closed form equations for option pricing cannot recreate the behavior observed in financial market data. This flaw in classical models leads to incorrect derivative pricing. Fang and Oosterlee (2008) introduced a method intended to recover the probability density functions in a quasi-analytical way. The authors demonstrated that, given the characteristic function of a continuously distributed random variable, it is feasible to estimate the probability density function pointwise using the Euler's identity and the Fourier-cosine series. The approach is effective and simple to adopt. The Fourier-cosine series is the central numerical method used along, throughout the paper.

Let $f: [0,\pi] \longrightarrow \mathbb{R}$ be an integrable function. Then the Fourier-cosine series of f is given by

$$f(\xi) = \frac{a_0}{2} + \sum_{j=1}^{\infty} a_j \cos(j\xi),\ \xi \in [0,\pi] \tag{11}$$

where

$$a_j = \frac{2}{\pi}\int_0^{\pi} f(\xi)\cos(j\xi)d\xi,\ j = 0,1,2,\ldots \tag{12}$$

For functions supported in any arbitrary interval [a,b], a change of variable ξ=π (x-a)/(b-a) is considered. Then, the Fourier-cosine series expansion of f in the interval [a,b] is

$$f(x) = \frac{a_0}{2} + \sum_{j=1}^{\infty} a_j \cos\left(j\pi\frac{x-a}{b-a}\right) \tag{13}$$

where

$$a_j = \frac{2}{b-a}\int_a^b f(x)\cos\left(j\pi\frac{x-a}{b-a}\right)dx,\ j = 0,1,2,\ldots \tag{14}$$

Let us assume that $f \in L^1(\mathbb{R})$. By the Euler's identity, the coefficients of the Fourier-cosine expansion of f are

$$\begin{aligned} a_j &= \frac{2}{b-a}\int_a^b f(x)\Re\left[\exp\left(ij\pi\frac{x-a}{b-a}\right)\right]dx \\ &= \frac{2}{b-a}\Re\left(\exp\left(-ij\pi\frac{a}{b-a}\right)\int_a^b f(x)\exp\left(ij\pi\frac{x}{b-a}\right)dx\right). \end{aligned} \tag{15}$$

Let X be a continuous random variable. If f, with domain in $\mathbb{R}$, is a probability density function of X, then

$$a_j \approx \frac{2}{b-a}\Re\left(\exp\left(-ij\pi\frac{a}{b-a}\right)\hat{f}\left(\frac{j\pi}{b-a}\right)\right) \triangleq A_j \tag{16}$$

where $\hat{f}$ is the characteristic function of X, that is

$$\hat{f}(u) = \int_{\mathbb{R}} \exp(ixu)f(x)dx \tag{17}$$

which approximates

$$\int_a^b \exp(ixu)f(x)dx \tag{18}$$

Therefore, the approximation of f is given by the following Fourier-cosine series

$$f(x) \approx \frac{A_0}{2} + \sum_{j=1}^{n} A_j \cos\left(j\pi\frac{x-a}{b-a}\right),\ x \in [a,b] \tag{19}$$

for a given n. In terms of acuity, the greater the n the better it is. Actually, as we shall see ahead, convergence is very fast and small numbers n will cope well with the aproximation.

The choices of the integration limits for the approximation be good were proposed in Fang and Oosterlee (2008) as follows:

$$a = c_1 - L\sqrt{c_2 + \sqrt{c_4}}\ b = c_1 + L\sqrt{c_2 + \sqrt{c_4}} \tag{20}$$

with L=10. The coefficients c_k are the k-th cumulant of x given by

$$c_k = \frac{1}{i^k}\frac{d^k}{du^k}h(u)\Big|_{u=0} \tag{21}$$

where the cumulant generating function is given by

$$h(u) = \ln\ \mathbb{E}\left[e^{iuX}\right]. \tag{22}$$

Remind that the domain of f, typically, is not [a,b]. This interval is now chosen in order to capture as much probability as possible from f.

Example 1. (Normal distribution) The normal distribution is a symmetric probability density function with mean μ and variance σ^2. Its characteristic function is given by

$$\hat{f}(u;\mu,\sigma) = e^{iu\mu-\frac{1}{2}\sigma^2u^2} \tag{23}$$

According to Equation (22), we have

$$h(u) = iu\mu - \frac{1}{2}\sigma^2u^2 \tag{24}$$

and $c_1 = \mu$, $c_2 = \sigma^2$ and $c_4 = 0$. In Figure 1 we show the approximation of the normal probability density function with parameters μ=10 and σ=2 as a function of n in the Fourier Series. We can note that the sum given by (19) converges fastly with few terms to the analytical probability density function given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{25}$$

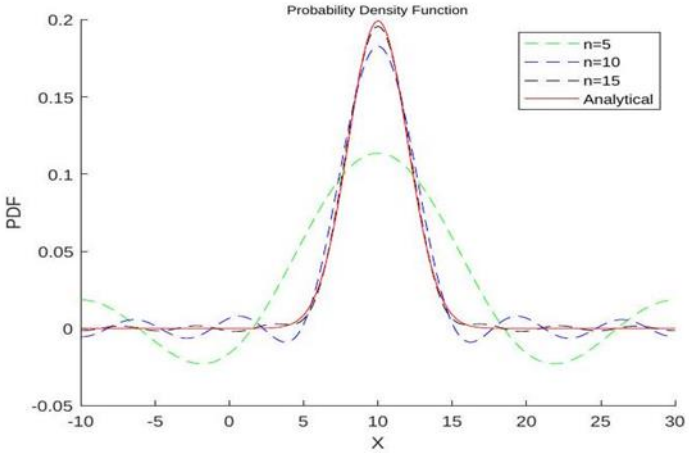


Figure 1. Convergence as a function of n.

2.4 The variance-gamma process

Following Madan et. al (1998), the variance-gamma process is obtained by evaluating Brownian motion with drift at a random time given by a gamma process. Let

$$b(t;\theta;\sigma) = \theta t + \sigma W(t), \tag{26}$$

where W(t) is a standard Brownian motion. The process $b(t;\theta,\sigma)$ is a Brownian motion with drift θ and volatility σ. The gamma process $\gamma(t;m,\beta)$ with mean rate m and variance rate β is the process of independent gamma increments over non-overlapping intervals of time (t,t+h). The density function of the increment $g = \gamma(t+h;m,\beta) - \gamma(t;m,\beta)$ is given by the gamma density function with mean mh and variance βh. Such process was first considered for stock prices by Clark (1973).

The variance-gamma process $X(t;\sigma,\beta,\mu)$, is defined in terms of the Brownian motion with drift b(t;μ,σ) and the gamma process with unit mean rate, $\gamma(t;1,\beta)$ as

$$X(t;\sigma,\beta,\mu) = b(\gamma(t;1,\beta);\mu;\sigma). \tag{27}$$

The variance-gamma process is obtained on evaluating Brownian motion at a time given by the gamma process. We can observe above that μ and β provide control over skew and kurtosis. More details can be found in the seminal paper Madan et. al (1998), in Zhu (2009) and in Tankov and Cont (2003).

2.5 Reinforcement learning

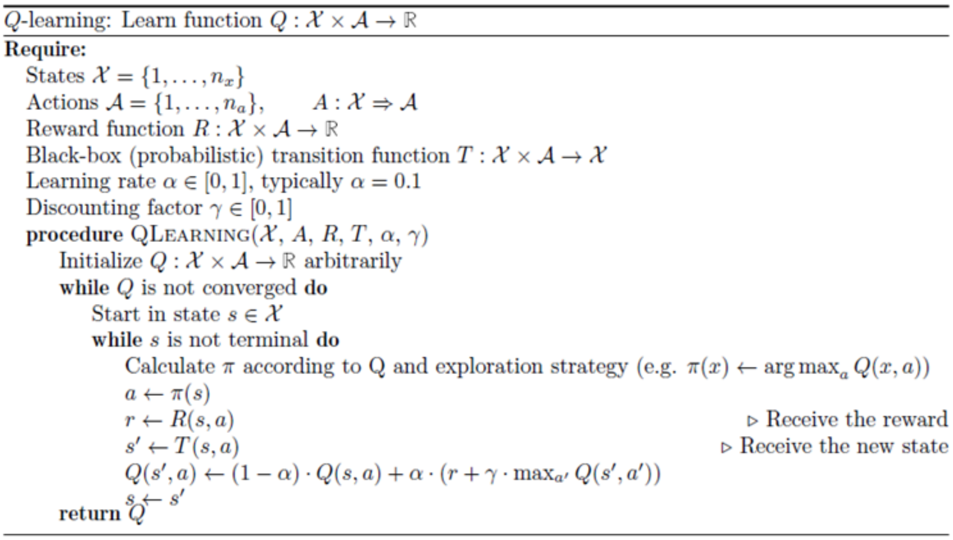
Reinforcement learning (RL) is a subfield of machine learning that focuses on training agents to make sequential decisions in an environment to maximize a reward signal (Kaelbling et al., 1996) or a cumulative reward. It has gained significant attention in recent years due to its ability to solve complex decision-making problems in various domains, such as autonomous driving (Sallab et al., 2017), lane-tracking control (Kalapos et al., 2021), and traffic signal control (Chu et al., 2020).

These applications leverage reinforcement learning algorithms to learn optimal policies for navigating complex environments and making real-time decisions. Similarly, in industrial autonomy, reinforcement learning has been used for part flow management in gas turbine maintenance (Compare et al. (2018)) and adaptive dispatching in semiconductor manufacturing systems (Sakr et al., 2021). These applications demonstrate the ability of reinforcement learning to optimize processes and improve efficiency in industrial settings.

Reinforcement learning has also found applications in finance and trading. It has been used for portfolio optimization (Joshi, 2022), where the goal is to allocate investments to maximize returns while managing risk. Additionally, reinforcement learning has been applied to algorithmic trading, where agents learn to make buy and sell decisions based on market conditions (Abdulhameed & Lupenko, 2022). These applications highlight the potential of reinforcement learning in the financial sector for making informed and profitable decisions.

This work employs Q-learning in order to hedge a call option in which the underlying asset path does not follow the Black-Scholes assumptions. Several research have been conducted to investigate the application of RL in option pricing. Du et al. (2020) offered a deep RL technique for option replication and hedging, for example. They developed a deep reinforcement learning agent to mimic option payoffs by dynamically modifying a portfolio of underlying assets. By maximizing the cumulative payoff, which was characterized as replication accuracy, the agent learned optimal trading strategies. Cao et al. (2019) proposed a deep hedging framework for dynamically hedging stock options in the Black-Scholes world using RL. They modeled the hedging problem as a Markov decision process and trained a deep RL agent to discover the best hedging strategy. Based on market conditions, the agent changed the hedge portfolio to reduce hedging costs and variation. Their research found that RL-based hedging techniques outperformed traditional approaches, especially when market frictions such as transaction costs were present.

The following is a step-by-step explanation of the Q-learning algorithm:



```
$Q$-learning: Learn function $Q : \mathcal{X} \times \mathcal{A} \to \mathbb{R}$
Require:
  States $\mathcal{X} = \{1, \ldots, n_x\}$
  Actions $\mathcal{A} = \{1, \ldots, n_a\}$,    $A : \mathcal{X} \Rightarrow \mathcal{A}$
  Reward function $R : \mathcal{X} \times \mathcal{A} \to \mathbb{R}$
  Black-box (probabilistic) transition function $T : \mathcal{X} \times \mathcal{A} \to \mathcal{X}$
  Learning rate $\alpha \in [0, 1]$, typically $\alpha = 0.1$
  Discounting factor $\gamma \in [0, 1]$
  procedure QLEARNING($\mathcal{X}$, $A$, $R$, $T$, $\alpha$, $\gamma$)
    Initialize $Q : \mathcal{X} \times \mathcal{A} \to \mathbb{R}$ arbitrarily
    while $Q$ is not converged do
      Start in state $s \in \mathcal{X}$
      while $s$ is not terminal do
        Calculate $\pi$ according to Q and exploration strategy (e.g. $\pi(x) \leftarrow \arg\max_a Q(x, a)$)
        $a \leftarrow \pi(s)$
        $r \leftarrow R(s, a)$                                  ▷ Receive the reward
        $s' \leftarrow T(s, a)$                                 ▷ Receive the new state
        $Q(s', a) \leftarrow (1 - \alpha) \cdot Q(s, a) + \alpha \cdot (r + \gamma \cdot \max_{a'} Q(s', a'))$
        $s \leftarrow s'$
    return $Q$
```

Figure 2. Q-learning algorithm. (Adapted from Sutton and Barto (2018))

3. Methodology

Following da Silva et. al (2019) and da Silva et. al (2020), with the focus on the interest rate derivatives market, let the continuous real valued random variable Z(t,T) be a function of the underlying source of risk - the interest rate process $\{r(s), s \in [t, T]\}$- experienced by an European call option maturing at time T. So, we may write Z(t,T)≡Z(t,$\{r(s), s \in [t, T]\}$). Let $f(\cdot | \, r(t))$ be the risk-neutral probability density function of Z(t,T) conditional to r(t) and g(Z(t,T)) be the discounted payoff function of the option. Then the price of this option at time t is

$$\begin{aligned} C(t,T) &= \mathbb{E}[g(Z(t,T)) \mid r(t)] \\ &= \int_{\mathbb{R}} g(w) f(w \mid r(t)) dw \end{aligned} \tag{28}$$

where $\mathbb{E}$ is the risk-neutral expected value. Truncating f in the interval [a,b] we have:

$$C(t,T) \approx \int_a^b g(w) f(w \mid r(t)) dw \tag{29}$$

The integral in (29) can be calculated by the COS method proposed by Fang and Oosterlee (2008). The COS method is an interesting, fast and accurate derivatives pricing method based on Fourier-cosine series. In what follows, we present the COS method and show how to use it to price options. Therefore, using f(x) as in (19) we have

$$C(t,T) \approx \frac{A_0}{2} \int_a^b g(x) \mathrm{d}x + \sum_{j=1}^n A_j \int_a^b g(x) \cos\left(j\pi \frac{x-a}{b-a}\right) \mathrm{d}x. \tag{30}$$

Hence, the series approximation of the option price is given by

$$C(t,T) \approx \frac{A_0 B_0}{2} + \sum_{j=1}^n A_j B_j, \tag{31}$$

where the A_k coefficients are given by (16) with the characteristic function (17) associated to the model and

$$B_j = \int_a^b g(x) \cos\left(j\pi \frac{x-a}{b-a}\right) \mathrm{d}x, \text{ for } j = 0,1,\dots,n, \tag{32}$$

associated to the payoff function.

3.1 The COS coefficients for the IDI call option

In order to employ the COS method to calculate the price of the IDI call option, it is necessary to find the B_j coefficients associated to the payoff of the IDI call option.

Theorem 3.1. The B_j coefficients for IDI call options are given by

$$B_0 = \int_{\ln(k)}^b (e^x - k) dx = k\ln(k) + (-b-1)k + \mathrm{e}^b, \tag{33}$$

and

$$\begin{aligned} B_j = &\int_{\ln(k)}^b (e^x - k)\cos\left(\frac{\pi j(x-a)}{b-a}\right) dx \\ = &(b-a)\frac{\left((b^2-2ab+a^2)k\sin\left(\frac{\pi j\ln(k)-\pi a j}{b-a}\right)+(\pi a-\pi b)jk\cos\left(\frac{\pi j\ln(k)-\pi a j}{b-a}\right)\right)}{\pi j(\pi^2 j^2+b^2-2ab+a^2)} + \cdots \\ &+(b-a)\frac{\left((-\pi^2 j^2-b^2+2ab-a^2)\sin(\pi j)k+\pi^2\mathrm{e}^b j^2\sin(\pi j)+(\pi b-\pi a)\mathrm{e}^b j\cos(\pi j)\right)}{\pi j(\pi^2 j^2+b^2-2ab+a^2)} \end{aligned} \tag{34}$$

Note that the B_j coefficients (33) and (34) are given by the Equation (32) using $g(x) = e^x - k$, where x stands for the variable that representes the random variable of the model, i.e. the logarithm of the present value of the forward ID index, and k the strike price of the contract.

3.2 Q-learning

Q-learning is a versatile algorithm that is frequently utilized in a variety of applications. According to Cao (2021), the policy update phase of the Q-learning approach incorporates a global optimization at each time step. Nonetheless, when the action space is continuous, as in the case of option hedging, this maximization becomes impractical.

Cao (2021) used a method called Deep Deterministic Policy Gradient (DDPG) to leverage Q-learning in conjunction with artificial neural networks. The Deep Deterministic Policy Gradient (DDPG) is a sophisticated reinforcement learning system that incorporates ideas from both Q-learning and policy gradient methods. It is specifically intended for handling high-dimensional, continuous action space issues, such as dynamic hedging. A complete explanation of the algorithm can be found in Cao (2021) and in Openai (2018).

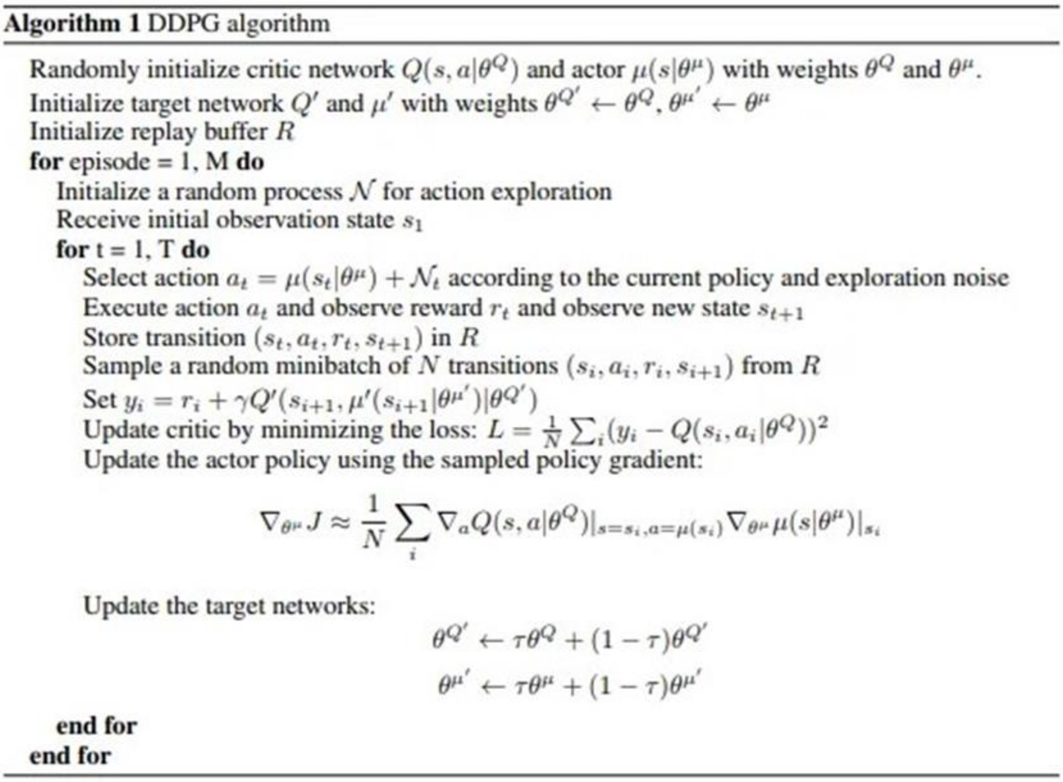
The agent's goal is to choose actions to maximize the expected cumulative reward:

$$\mathbb{E}[G_t] = \mathbb{E}[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots + \gamma^{T-1} R_T], \tag{35}$$

where the constant $\gamma \in (0,1)$ is referred to as the discount fator and T is the maturity of the option. The reward is given by the profit and loss formulation of Cao (2021):

$$R_{i+1} = H_i(S_{i+1} - S_i) + V_{i+1} - V_i - \kappa|H_{i+1} - H_i|S_{i+1}], \tag{36}$$

where S_i is the asset price, at time i, H_i is the holding quantity, κ the trading cost and V_i is the value of the option. We benefited from DDPG agent implemented in the reinforcement learning toolbox of matlab. The following algorithm was proposed in Lillicrap et al. (2015):



```
Algorithm 1 DDPG algorithm
  Randomly initialize critic network $Q(s, a|\theta^Q)$ and actor $\mu(s|\theta^\mu)$ with weights $\theta^Q$ and $\theta^\mu$.
  Initialize target network $Q'$ and $\mu'$ with weights $\theta^{Q'} \leftarrow \theta^Q, \theta^{\mu'} \leftarrow \theta^\mu$
  Initialize replay buffer $R$
  for episode = 1, M do
    Initialize a random process $\mathcal{N}$ for action exploration
    Receive initial observation state $s_1$
    for t = 1, T do
      Select action $a_t = \mu(s_t|\theta^\mu) + \mathcal{N}_t$ according to the current policy and exploration noise
      Execute action $a_t$ and observe reward $r_t$ and observe new state $s_{t+1}$
      Store transition $(s_t, a_t, r_t, s_{t+1})$ in $R$
      Sample a random minibatch of $N$ transitions $(s_i, a_i, r_i, s_{i+1})$ from $R$
      Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta^{\mu'})|\theta^{Q'})$
      Update critic by minimizing the loss: $L = \frac{1}{N}\sum_i (y_i - Q(s_i, a_i|\theta^Q))^2$
      Update the actor policy using the sampled policy gradient:

        $\nabla_{\theta^\mu} J \approx \frac{1}{N}\sum_i \nabla_a Q(s, a|\theta^Q)|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^\mu}\mu(s|\theta^\mu)|_{s_i}$

      Update the target networks:
        $\theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta^{Q'}$
        $\theta^{\mu'} \leftarrow \tau\theta^\mu + (1 - \tau)\theta^{\mu'}$

    end for
  end for
```

Figure 3. DDPG algorithm. (Lillicrap et al. (2015))

4. Results

Interest rate options behave in the opposite way to equities. Typically, investors want insurance against an increase in interest rates. As a result, the higher the strike, the greater the demand for options, implying higher costs.

When interest rates rise, the value of fixed-income products like bonds falls. This occurs as newer bonds with higher coupon rates enter the market, making older bonds with lower coupon rates less appealing. Investors who hold these bonds may suffer losses if bond prices fall. Investors use interest rate options to offset this risk (Goodman and Fabozzi (2002)).

Investors who purchase interest rate call options gain the right to purchase an underlying instrument (such as a bond) at a preset strike price even if interest rates climb. This provides some insurance since the option increases in value when the underlying security's price decreases owing to rising interest rates. Thus, the higher the strike price specified in the option contract, the greater the amount of protection provided, resulting in more demand and, as a result, higher option pricing.

The volatility smile of IDI options, unlike stock options, has a forward skew structure. Deep out-of-the-money calls and deep in-the-money puts are richer than which is precluded by the Black-76 model (Black (1976)), the standard model used by the Brazilian exchange.

Figure 2 shows a typically risk-neutral probability density function implied from a 6-month IDI call option. We use the numerical second derivative of the option price with respect to the strike as an approximation of the density function, as proposed by Breeden and Litzenberger (1978). In the particular case of IDI option, from equation (4) we have

$$e^x \frac{\partial^2 C(t,T)}{\partial K^2} = f(x), \tag{37}$$

where $x = \int_t^T r(s)ds$. Note that we chose to recover the density of the integral of the interest rate along the remaining time of the option. We could work with the density of the index y(t) itself, which would result in

$$\frac{\partial^2 C(t,T)}{\partial K^2} = f(y). \tag{38}$$

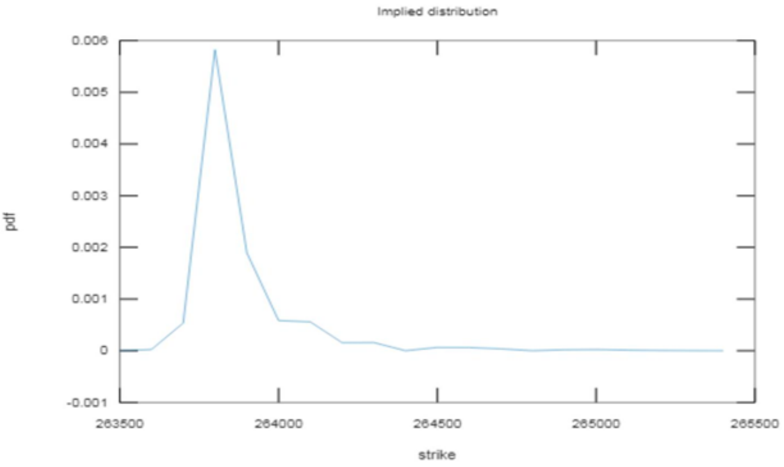


Figure 4. Implied probability distribution from IDI option prices.

We note in this example that the implied distribution is leptokurtic and positively skewed. Right long-tailed implied distributions are often found in market prices of IDI options of various maturities.

The models presented up to now are poor concerning the ability to recover the market prices features. In this paper we seek a model capable to reproduce the behavior of the empirical distribution in conjunction with an efficient numerical method.

4.1 The Black model

In order to calculate the IDI option price with the COS method, according to Fang and Oosterlee (2008), it is necessary to express the characteristic function of the underlying model as a closed form expression.

We assume a geometric Brownian motion with zero drift to modeling the present value of the forward ID index $\bar{y}(t)$ as follows:

$$d\bar{y}(t) = \sigma\bar{y}(t)dW(t). \tag{39}$$

The Black-76 model is the reference formula for the market. The model implies a lognormal probability density function for $\bar{y}(t)$, which is not able to capture the stylized facts of the market data.

We analyze the IDI option prices given by the COS method. Firstly, we assume that the present value of the forward ID index evolves according to the Black-76 model. In this simple case, Carreira and Brostowicz, (2016). provided a closed-form solution to the IDI call price shown in (9), which allows us to investigate the error of the COS method. To employ the COS method, we need to insert into the equation (16) the conditional characteristic function of $\bar{y}(T)$, which is given by

$$\hat{f}(\bar{y}(t), t, u) = e^{iu\left[\log(\bar{y}(t)) - \frac{\sigma^2(T-t)}{2}\right] - \frac{u^2\sigma^2(T-t)}{2}}. \tag{40}$$

Cumulant functions (21) for the Black-76 model can be found in Kienitz and Wetterau (2012) and Fang and Oosterlee (2008). Characteristic functions of financial models are broadly discussed and implemented in Kienitz and Wetterau (2012).

In this exercise, we work with σ=0.06%. The time to maturity of the option is 54 business days and its strike price is 262500. The IDI spot y(t) is 260321.410. Figure 3 shows the prices of this option as a function of the strike calculated by the COS method and by the closed-form formula. Note that the prices are visually indistinguishable.

The COS method produces 25 prices for N=32 terms in the summation (30) in less than a half of a second in a standard personal computer.

4.2 Black model with Lévy jumps case: the variance-gamma process

In order to embody the stylized facts of the real world on the implied distribution of the IDI option, we use a version of the Black-76 model enhanced with a Lévy process as

$$d\bar{y}(t) = \bar{y}(t)dL(t) \tag{41}$$

where the variance-gamma process L is a pure jump Lévy process defined in terms of the Brownian motion with drift $\theta t + \sigma W(t)$, subordinated by a Gamma process γ(t,α,β), as described in Subsection 2.4. The latter therefore is an increasing Lévy process. The so called VG process is discussed in Tankov and Cont (2003), Zhu (2009) and Kienitz and Wetterau (2012).

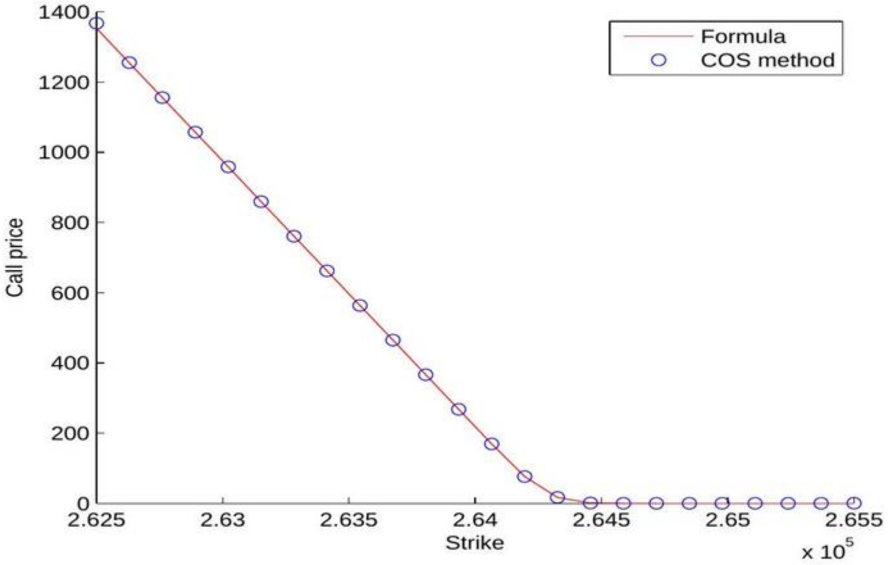


Figure 5. Black-76 prices

We have that α=1 is the mean rate and β is the variance rate of the Gamma process. Hence, substituting t by $\gamma(t,\alpha,\beta)$ in θt+σW(t) gives us the Variance-Gamma process

$$L(t) = \mu \cdot \gamma(t,\alpha,\beta) + \sigma W(\gamma(t,\alpha,\beta)) \tag{42}$$

The corresponding conditional characteristic function of the variance-gamma model for time T to be inserted in (16) is

$$\hat{f}(u;\mu,\beta,\sigma) = e^{iu\ln(\bar{y}(t))}\left(1 - iu\mu\beta + \frac{1}{2}\sigma^2u^2\beta\right)^{-\frac{(T-t)}{\beta}} \tag{43}$$

Details about Lévy process and its characteristic function can be found in Tankov and Cont (2003) and Zhu (2019). Cumulant functions (21) for the variance-gamma model can be found in Kienitz and Wetterau (2012) and Fang and Oosterlee (2008).

In Figure 4 we show a range of the IDI option prices given by (39), where L(t) is the variance-gamma process with parameters μ, β, σ. We compare the call option prices of the Black-76 model with those of the enhanced Black-76 with variancegamma process. We replicate the parameters of the above example and use β=2.5.

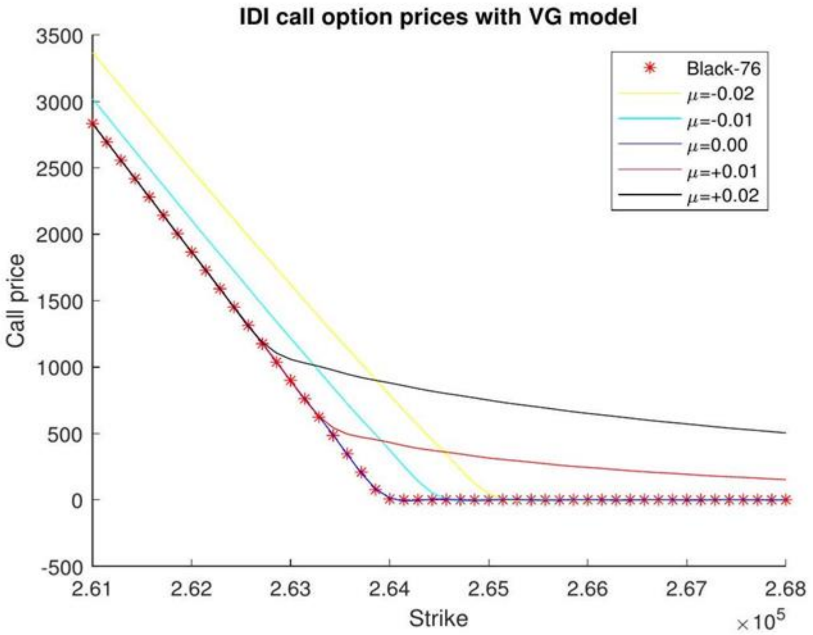


Figure 6. Comparison of IDI Call prices with Black-76 and Black-variance-gamma processes

By varying the parameter μ betwenn -0.02 and 0.02, we can observe the power of the VG process. We note that for negatives values of μ, the in-the-money IDI options prices are much richer than those of the Black-76 model. Conversely, for positive values of μ, out-of-the-money IDI options prices are much richer than those given by the Black-76 model. All four cases produce higher prices than those of the Black-76 model for at-the-money options. Note that, for μ=0, the VG and the Black-76 model prices match for this set of parameters.

The preceding experiment demonstrated that the VG process can create excess kurtosis and a right fat-tail - for positive μ as we seek - as shown in IDI option market data. For example, the potential of positive leaps implies significantly higher-priced options, implying that the likelihood of the option expiring in-the-money was increased. Figure 7 depicts typical volatility smiles formed in the IDI option market produced variance-gamma process. It is important to recover that, the market practice is to find the implied volatility by numerically inverting the Black-76 and finding σ, which is constant. It is worth noting that the model is capable of capturing the properties of options found in market data, such as a growing volatility curve.

Through a comparison of the Black-76 model and the upgraded Black-76 model utilizing the variance-gamma process, the results reported in this section provide useful insights into the behavior of IDI option pricing. The empirical investigation provides light on the impact of altering the parameter μ inside the VG process, providing a thorough grasp of its impact on volatility curve, and consequently on option pricing, at various moneyness levels. We observe that for this set of parameters, negative μ produces the classic volatility smile seen in stock markets. For instance, negative μ it is what was found for stock options data in the seminal paper of Madan et. al (1998).

The previous experiment demonstrates the VG process's ability to reproduce the heavy-tail features reported in IDI option market data. The model's capacity to represent real-world phenomena is highlighted by the development of excess kurtosis and a right fat-tail, controled with β and μ, which are hallmarks of interest rate market behaviors characterised by severe occurrences. The appearance of the volatility smile in Figure 5, created with positives μ values, reflects the typical phenomenon found in option markets, confirming the VG process's practical use.

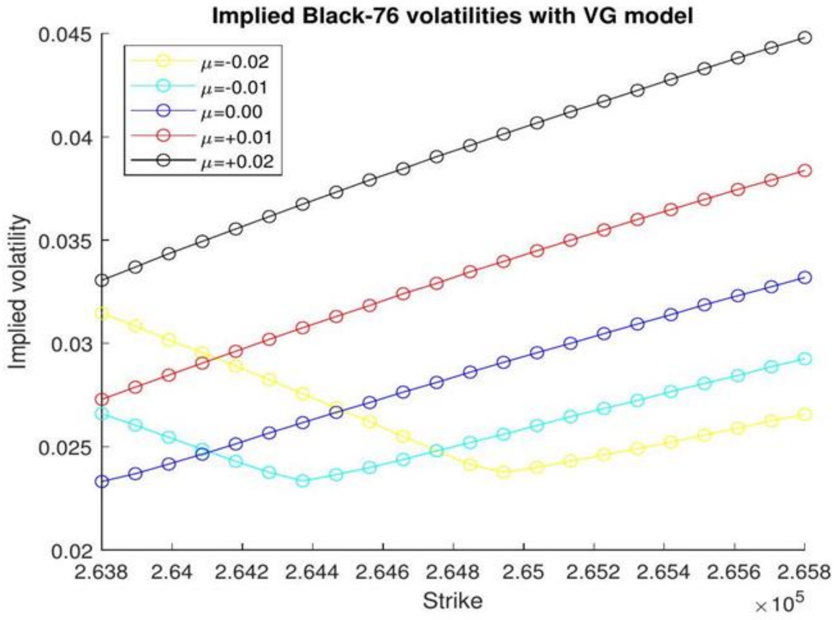


Figure 7. VG model implied volatility smile

In addition, does the computational technique described here confirm the VG process's effectiveness in catching heavy-tail activity, but it also provides a more comprehensive understanding of enhancing option pricing accuracy using the COS method. This work adds significantly to our understanding of financial market dynamics and expands the toolset accessible to practitioners in their quest of better risk management and pricing precision.

Considering that the preceding experiments provide empirical evidence of the VG model's ability to replicate market index behavior, it becomes feasible to employ this model for investigating the impact of reinforcement learning-based hedging.

4.3 Hedging

The set of experiments detailed in what follows delves into the dynamic world of financial options hedging. In an environment where market conditions are far from idealized, transaction costs play a significant role, and index movements follow complex patterns like the variance-gamma process, it becomes essential to explore innovative and adaptive hedging strategies. These experiments compare the traditional Black-76 delta hedging method with a reinforcement learning (RL) approach, trained using the Q-learning algorithm. The primary aim is to assess how RL can impact the effectiveness and cost-efficiency of hedging under varying scenarios, ultimately shedding light on its potential to revolutionize hedging strategies in the real-world context of, at least, interest rate derivatives.

We consider first that the index evolves as the Black-76 model. We consider the spot price = 100 and at-the-money call option. The maturity is equal to 0.25 years and σ = 0.2. The annual interest rate equals 0.1. From now on we consider discretely daily hedging. The Q-learning algorithm was trained by using 5000 trials. Using the Black-76 formula, the hedging quantity is given by

$$\Delta = e^{-rT}N(d_1). \tag{44}$$

- By considering no costs of transaction, the hedging results are as follows:

  The Black-76 delta hedging performed with Equation (44) resulted in an average hedge cost of 1.2% of the option price. The RL hedging resulted in an average hedge cost of -1% of the option price. The standard deviations are 39% and 58%, respectively. Figure 8 illustrates the results.

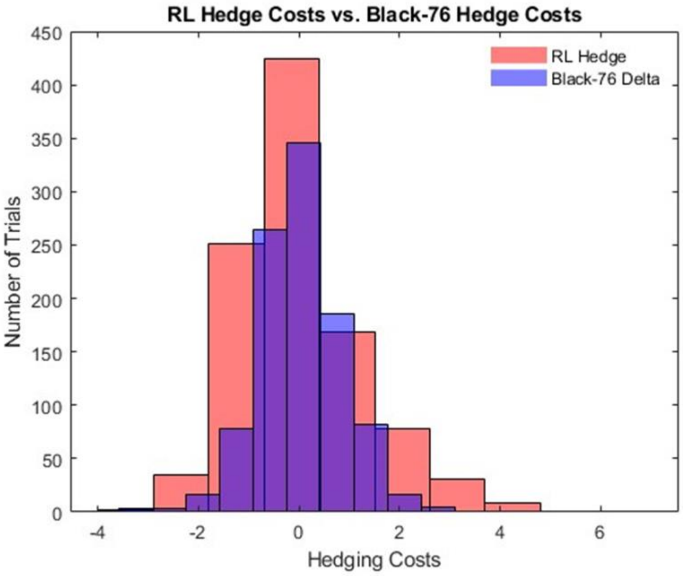


Figure 8. Hedging costs (RL vs Black-76 with no transaction cost)

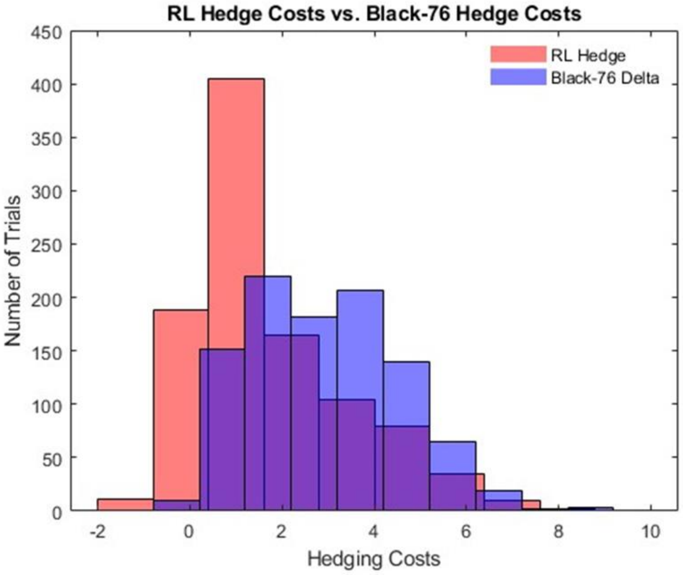


Figure 9. Hedging costs (RL vs Black-76 with 1% transaction cost)

- By considering the costs of the transaction of 1%, the hedging results are as follows:

  The Black-76 delta hedging performed with Equation (44) resulted in an average hedge cost of 141% of the option price. The RL hedging resulted in an average hedge cost of 82% of the option price. The standard deviations are 76% and 81%, respectively. The RL hedging reduces the hedging cost by 31%. Figure 9 illustrates the results.

- By considering the costs of the transaction of 1% and that the index evolves as a variance-gamma process, the hedging results are as follows:

  The Black-76 delta hedging performed with Equation (44) resulted in an average hedge cost of 112% of the option price. The RL hedging resulted in an average hedge cost of 79% of the option price. The standard deviations are 93% and 106%, respectively. The RL hedging reduces the hedging cost by 30%. Figure 10 illustrates the results:

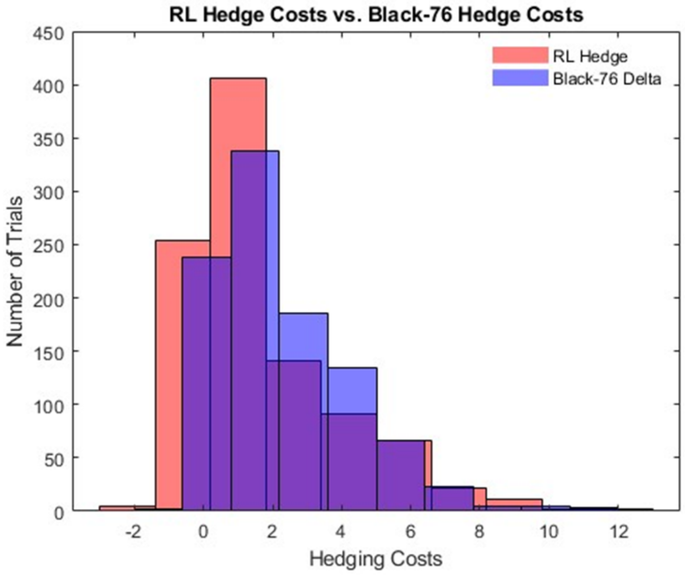


Figure 10. Hedging costs (RL vs Black-76 with 1% transaction cost and VG paths)

- In Figures 11, 12 and 13 we show the variance-gamma process sample path for y(t) given by the Equation (41), the corresponding option prices and the implied volatilities dynamics, respectively. Note that the implied volatility, as expected, is not constant over time. The variance-gamma process can indeed mimic market conditions.

The presented results offer valuable insights into the effectiveness of RL hedging in different scenarios. When transaction costs are introduced, the advantages of RL hedging become even more evident. These findings underscore the cost-saving capabilities of RL in realistic trading scenarios.

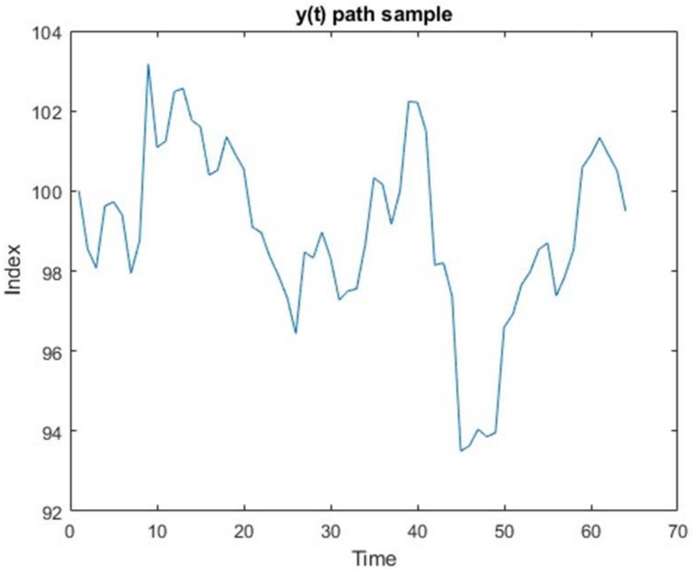


Figure 11. Variance-gamma process index simulated path

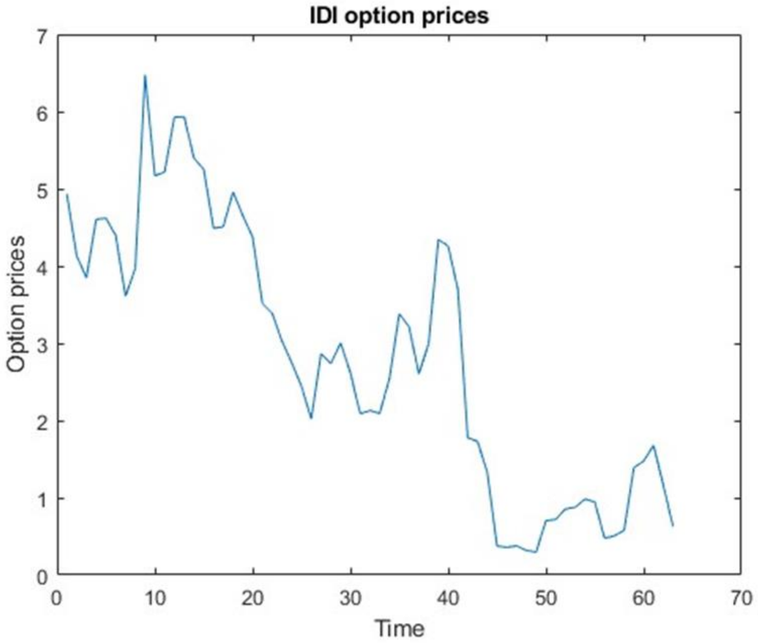


Figure 12. Variance-gamma option prices simulated path

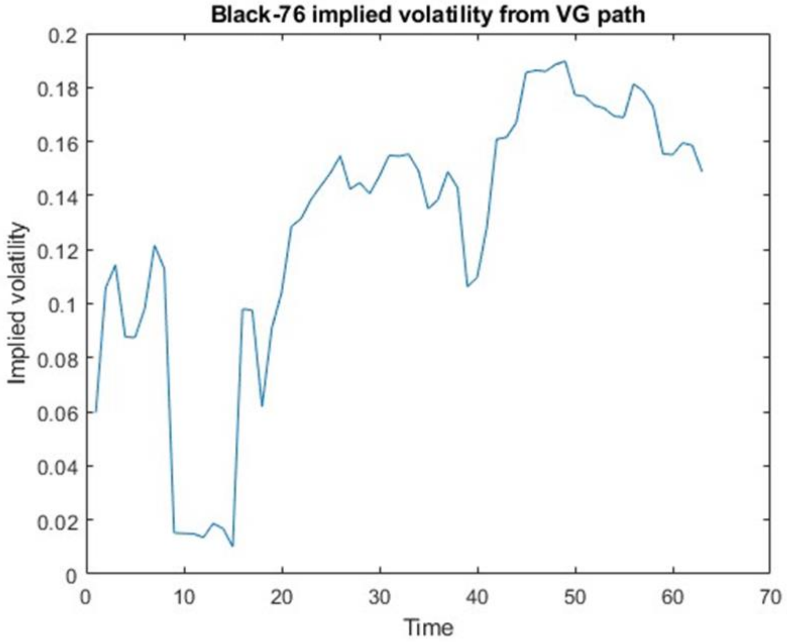


Figure 13. Black-76 implied volatilities simulated path generated from VG option prices

5. Conclusion

This study proposed a thorough statistical model targeted at capturing the peculiar heavy-tail behavior visible in option pricing in the Interbank Deposit Rate Index (IDI) market. The paper begins by outlining the mathematical foundations of IDI options as well as the traditional pricing methodology used by market participants to infer implied volatility. Following that, the Fourier cosine series approach (COS) is described in detail, with numerical computations of the probability density function conducted. The inquiry highlights the potential benefits of the COS technique in dealing with heavy-tail behaviors through a comparative examination of option prices obtained from both the Black-76 formula and the COS method.

The introduction of the variance-gamma (VG) process, a unique statistical model intended to adequately capture the observed heavy tail features, is important to the research. The empirical results confirm the VG process's potential to create excess kurtosis and a right fat-tail, both of which are characteristics of IDI option market data. The article explains the subtle impact of multiple factors within the VG process using illustrative data, emphasizing the influence of μ on option prices at different moneyness levels. The VG approach exhibits its ability to generate much richer in-the-money or out-of-the-money option prices, providing a viable path for improved pricing precision.

Furthermore, the study underscores the VG process's practical usefulness by demonstrating the creation of the well-known volatility smile - a hallmark of financial markets - in the IDI option market environment. Notably, the market practice of assessing implied volatility by numerical inversion of the Black-76 model is highlighted, emphasizing the research findings' practical importance for market players.

Finally, by considering transaction costs, alternative index dynamics, and the dynamic nature of implied volatilities, this research uncovers the transformative power of reinforcement learning in the realm of financial option hedging. The study introduced reinforcement learning hedging and extended its investigation to a scenario where the index follows a variance-gamma process, showcasing that reinforcement learning hedging remains more cost-effective compared to the Black-76 model, even under these alternative market conditions. These results illustrate RL's adaptability to different market dynamics, opening up a wide range of possibilities for conducting further experiments with new financial models and products.

Acknowledgments

This work was supported by CNPq under the grant 140840/2015-0.

Authors contributions

The variance-gamma model was intended to be used to characterize the IDI index value, and Drs. da Silva and Baczynski were in charge of conceptualizing the mathematical pricing strategy for IDI options. Dr. da Silva implemented the pricing model using Fourier series. Additionally, Dr. da Silva and Dr. Mello jointly implemented the discrete hedging problem using reinforcement learning, while Dr. Baczynski contributed to the analysis of the obtained results. All authors critically reviewed and approved the final manuscript prior to submission.

Funding

Not applicable.

Competing interests

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.

Informed consent

Obtained.

Ethics approval

The Publication Ethics Committee of the Redfame Publishing.

The journal’s policies adhere to the Core Practices established by the Committee on Publication Ethics (COPE).

Provenance and peer review

Not commissioned; externally double-blind peer reviewed.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author.

Data sharing statement

No additional data are available.

Open access

Copyrights

References

Abdulhameed, S. A. and Lupenko, S. (2022). Potentials of reinforcement learning in contemporary scenarios. Scientific Journal of the Ternopil National Technical University, 2(106), 92-100.

Almeida, C. and Vicente, J. V. M. (2012). Term structure movements implicit in Asian option prices. Quantitative Finance, 12(1):119-134.

Almeida, L. A., Yoshino, J., and Schirmer, P. P. S. (2003). Derivativos de renda-fixa no Brasil: Modelo de Hull-White. Pesquisa e Planejamento Econômico, 33:299-333.

Barbachan, J. S. F. and Ornelas, J. R. H. (2003). Apreçamento de opções de IDI usando o modelo CIR. Estudos Econômicos, 33(2):287-323.

Barbedo, C. H., Vicente, J. V. M., and Lion, O. B. (2010). Pricing Asian interest rate options with a three-factor HJM model. Revista Brasileira de Finanças, 8(1):9-23.

Black, F. (1976). The pricing of commodity contracts. Journal of Financial Economics, 3(1-2):167-179.

Bouziane, M. (2008). Pricing interest-rate derivatives: a Fourier-transform based approach. Springer, Berlin.

Breeden, D. and Litzenberger., R. (1978). State contingent prices implicit in option prices. Journal of Business, 51:621-651.

Brigo, D. and Mercurio, F. (2006). Interest Rate Models - Theory and Practice. Springer Finance. Springer, Berlin.

Cao, J. et al. Deep hedging of derivatives using reinforcement learning. SSRN Electronic Journal, 2019.

Cao Jay, C. J. H. J. P. Z. Deep hedging of derivatives using reinforcement learning. The Journal of Financial Data Science, v. 3, n. 1, p. 10-27, 2021.

Carreira, M. and Brostowicz, R. (2016). Brazilian Derivatives and Securities: Pricing and Risk Management of FX and Interest-Rate Portfolios for Local and Global Markets. Palgrave Macmillan UK.

Compare, M., Bellani, L., Cobelli, E., & Zio, E. (2018). Reinforcement learning-based flow management of gas turbine parts under stochastic failures. The International Journal of Advanced Manufacturing Technology, 99(9-12), 2981-2992.

Chu, T., Wang, J., Codecà, L., & Li, Z. (2020). Multi-agent deep reinforcement learning for large-scale traffic signal control. IEEE Transactions on Intelligent Transportation Systems, 21(3), 1086-1095.

Clark, P. (1973). A subordinated stochastic process model with finite variance for speculative prices. Econometrica, 41:135-156.

Cox, J. C., Ingersoll, J., and Ross, S. (1986). A theory of the term structure of interest rates. Econometrica, 53:385-407.

da Silva, A. J., Baczynski, J., and Bragança, J. F. S. (2019). Path-dependent interest rate option pricing with jumps and stochastic intensities. Lecture Notes in Computer Science, 11540:710-716.

da Silva, A. J., Baczynski, J., and Vicente, J. V. M. (2016). A new finite difference method for pricing and hedging fixed income derivatives: Comparative analysis and the case of an Asian option. Journal of Computational and Applied Mathematics, 297:98-116.

da Silva, A. J., Baczynski, J., and Vicente, J. V. M. (2020). Efficient solutions for pricing and hedging interest rate asian options. Working Paper Series - Banco Central do Brasil, (513).

da Silva, A. J., Baczynski, J., and Vicente, J. V. M. (2023). Recovering probability functions with fourier series. Pesquisa Operacional, 43:1-18.

Dandapani, K. (2017). Electronic finance – recent developments, Managerial Finance, Vol. 43 No. 5, pp. 614-626.

De Domenico, F., Livan, G., Montagna, G., and Nicrosini, O. (2023). Modeling and simulation of financial returns under non-gaussian distributions. Physica A: Statistical Mechanics and its Applications, 622:128886.

Du, J. et al. Deep reinforcement learning for option replication and hedging. The Journal of Financial Data Science, 2020.

Duffie, D. (2008). Financial Modeling with Affine Processes. Stanford University and University of Lausanne.

Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81:637-54.

Fabozzi, F. J. (2000). Mercados, Análise e Estratégias de Bônus. Qualitymark, Rio de Janeiro, 1a edition.

Fang, F. and Oosterlee, C. W. (2008). A novel pricing method for European options based on Fourier-cosine series expansions. SIAM Journal on Scientific Computing, 31(2):826-848.

Gafrej, O. (2023). Predicting customer deposits with machine learning algorithms: evidence from Tunisia, Managerial Finance, Vol. ahead-of-print No. ahead-of-print.

Gatheral, J. (2006). The volatility surface: a practitioner's guide. The Wiley Finance Series. Wiley.

Genaro, A. D. and Avellaneda, M. (2018). Pricing interest rate derivatives under monetary policy changes. International Journal of Theoretical and Applied Finance, 21(6):1850037.

Glasserman, P. (2004). Monte Carlo Methods in Financial Engineering. Applications of mathematics: stochastic modelling and applied probability. Springer.

Goodman, L. and Fabozzi, F. (2002). Collateralized Debt Obligations: Structures and Analysis. Frank J. Fabozzi Series. Wiley.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. Econometrica, 60(1):77-105.

Hull, J. and White, A. (1993). One-factor interest rate models and the valuation of interest-rate derivatives securities. Journal of Financial and Quantitative Analysis, 28(2):235-253.

Hull, J. C. (2009). Options, Futures and Others Derivatives. Pearson Prentice Hall, 7th edition.

Jondeau, E., Poon, S.-H., and Rockinger, M. (2007). Financial Modeling Under Non-Gaussian Distributions. Springer, London.

Joshi, D. (2022). Portfolio optimization using reinforcement learning. Interantional Journal of Scientific Research in Engineering and Management, 06(10).

Junior, A. F., Grecco, F., Lauro, C., Francisco, G., Rosenfeld, R., and Oliveira, R. (2003). Application of Hull-White model to Brazilian IDI options. In Annals of Brazilian Finance Meeting.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: a survey. Journal of Artificial Intelligence Research, 4, 237-285.

Kalapos, A., Gór, C., Moni, R., & Harmati, I. (2021). Vision-based reinforcement learning for lane-tracking control. Acta Imeko, 10(3), 7.

Khalaf, L., Leccadito, A., and Urga, G. (2021). Multilevel and Tail Risk Management*. Journal of Financial Econometrics, 20(5):839-874.

Kienitz, J. and Wetterau, D. (2012). Financial Modelling: Theory, Implementation and Practice with MATLAB Source. The Wiley Finance Series. Wiley.

Lillicrap, T. P. et al. (2015) Continuous control with deep reinforcement learning, CoRR abs/1509.02971.

Madan, D. B., Carr, P. P., and Chang, E. C. (1998). The variance gamma process and option pricing. European Finance Review, 2:79-105.

Marshall, J. F. and Bansal, V. K. (1991). Financial Engineering: a complete guide to financial innovation. New York Institute of Finance.

Martellini, L., Priaulet, P., and Priaulet, S. (2003). Fixed-income securities. John Wiley & Sons, England.

Nain, I. and Rajan, S. (2023), Algorithms for better decision-making: a qualitative study exploring the landscape of robo-advisors in India, Managerial Finance, Vol. ahead-of-print No. ahead-of-print.

Neftci, S. (2000). An Introduction to the Mathematics of Financial Derivatives. Elsevier, New York, 2nd edition.

Núñez - Letamendia, L. (2002), Trading systems designed by genetic algorithms, Managerial Finance, Vol. 28 No. 8, pp. 87-106.

Openai. Deep Deterministic Policy Gradient. 2018. < https://spinningup.openai.com/en/ latest/algorithms/ddpg.html>.

Ornelas, J. R. H. and Takami, M. Y. (2011). Recovering risk-neutral densities from brazilian interest rate options. Brazilian Review of Finance, 9(1):9-26.

Sakr, A. H., AboElHassan, A., Yacout, S., & Bassetto, S. (2021). Simulation and deep reinforcement learning for adaptive dispatching in semiconductor manufacturing systems. Journal of Intelligent Manufacturing, 34(3), 1311-1324.

Sallab, A. E., Abdou, M., Pérot, E., & Yogamani, S. (2017). Deep reinforcement learning framework for autonomous driving. Electronic Imaging, 29(19), 70-76.

Smithson, C. (1998). Managing Financial Risk: A Guide to Derivative Products, Financial Engineering, and Value Maximization. McGraw-Hill, 3rd edition.

Sutton, R., and Barto, A. (2018). Reinforcement Learning: An Introduction. The MIT Press, Second edition.

Tankov, P. and Cont, R. (2003). Financial Modelling with Jump Processes. Financial Mathematics Series. Chapman & Hall/CRC, Florida.

Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5:177-188.

Vellekoop, M. and Nieuwenhuis, H. (2007). On option pricing models in the presence of heavy tails. Quantitative Finance, 7(5):563-573.

Vieira, C. and Pereira, P. (2000). Closed form formula for the price of the options on the 1 day brazilian interfinancial deposits index. In Annals of the XXII Meeting of the Brazilian Econometric Society, volume 2, Campinas, Brazil.

Wilmott, P. (2006). Paul Wilmott on Quantitative Finance. John Wiley & Sons, Chichester, 2th edition.

Zhu, J. (2009). Applications of Fourier Transform to Smile Modeling: Theory and Implementation. Springer Finance. Springer Berlin Heidelberg.