Pricing American Options through Reinforcement Learning

By

YOUSSEF AIT HADDOU

195460

BACHELOR OF SCIENCE IN MATHEMATICS WITH HONOURS

DEPARTMENT OF MATHEMATICS
FACULTY OF SCIENCE
UNIVERSITI PUTRA MALAYSIA

FIRST SEMESTER 2020/2021

ABSTRACT

PRICING AMERICAN OPTIONS THROUGH REINFORCEMENT LEARNING

By

YOUSSEF AIT HADDOU

Supervisor: Dr. Chen Chuei Yee

Department: Mathematics

Faculty: Science

Options have become one of the most important instruments in finance, which allows the buyers to buy or sell an underlying asset at a fixed price. In this study, we discuss American style options, which allows the holder of the option to exercise it at any time up to the expiration date. Since the Longstaff-Schwartz method is considered the most widely used option pricing method, we want to examine if reinforcement learning methods, mainly least squares policy iteration and fitted Q iteration, can provide better pricing of auto-callable structured products than the classical Longstaff-Schwartz method. We work conduct this study withing Monte-Carlo framework, where we implement mathematical models, namely Black-Scholes and Heston models to simulate the underlying price dynamics and generate different possible paths of the price. Then, we apply reinforcement learning methods to price auto-callable structured products, where we calculate the fair value and the corresponding Delta risk of this option. Finally, we conduct a realistic back-test of a Delta hedged portfolio using the historical data of Apple Inc. (AAPL) in the duration of 2 years. Our empirical results show that the reinforcement learning methods provide a significant 330% risk reduction.

ACKNOWLEDGMENTS

This final year project, titled Pricing American Options through Reinforcement Learning, was prepared for Department of Mathematics, University Putra Malaysia (UPM) to complete the undergraduate program that leads to the degree of Bachelor of Science in Mathematics.

First of all, I would like to thank my family for their continuous support that has been shown to me throughout my time in the university, especially during the COVID-19 pandemic. I would like to express my to my advisor, Dr. Chen for her incredible guidance and irreplaceable feedback with patience and unlimited motivation during this research. I wish to express my appreciation to my brother Brahim for providing me with indispensable information and hands-on knowledge that allowed to me to successfully conduct this study. Finally, the completion of this project required a lot of guidance and assistance from many people, thanks go out to all my professors and my fellow undergraduate students at UPM for making my experience in the university enjoyable and fruitful.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Option valuation has become one of the most important topics in the theory of finance. It provides the theoretical fair value of the option which traders should incorporate in their analysis and strategies. Options are financial derivatives that give holders the right, but not the obligation, to buy or sell an underlying asset at a fixed price known as a strike price or an exercise price, and at or before the expiration date of the option. These financial instruments come in two basic forms:

- Call option: also referred to as calls, this contract gives the holder the right to purchase the security at the strike price before the maturity date.
- Put option: also referred to as puts, this contract gives the holder the right to sell the security at the strike price before the maturity date

Let S_T be the underlying price at the maturity date T, X the strike price and P the option value, and the payoff of a call option is defined as $max(S_T - X, 0)$ while the payoff of a put option is $max(X - S_T, 0)$. For instance, the put option buyer will receive $max(X - S_T, 0)$ and his profit is $max(X - S_T, 0) - P$ while the seller will receive an amount of $-max(X - S_T, 0)$ and the profit is $P - max(X - S_T, 0)$. In the case of $S_T < X$ the option is said to finish in the money (ITM), the buyer of the option will exercise the right to sell this option at the strike price X and gets a payoff of of $X - S_T > 0$. However, if $S_T > X$ which implies that $X - S_T < 0$, the option is said to finish out of the money (OTM), thus the holder of the option will not exercise the option to avoid a loss of $X - S_T$ and the payoff will be zero. The same applies for call options with respect to the call option payoff which is $max(S_T - X, 0)$. Figure 1.1 illustrates the payoffs and profits of buying or selling call and put options.

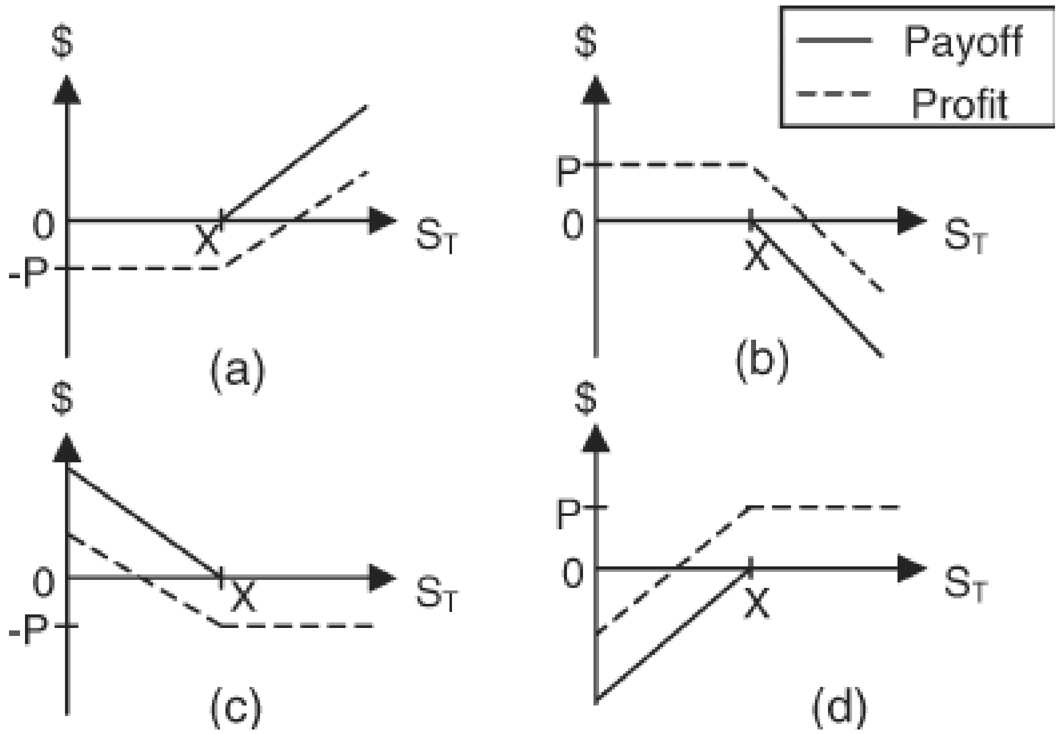


Figure 1.1: payoff and profit for a call buyer (a) and seller (b), and the payoff and profit for a put buyer (c) and seller (d) by Enke and Amornwattana (2008).

There are two major types of options, namely the European and American. The European option gives the investors the chance to exercise the contract at a fixed date, but an an early exercise is not allowed. Black and Scholes (1973) showed that this style of options satisfies a differential equation which became the Black-Scholes equation:

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{1.1}$$

where

$$
\begin{aligned}
V &: \text{ Price of the option,} \\
r &: \text{ Risk-free interest rate,} \\
\sigma &: \text{ Volatility of the underlying,} \\
S &: \text{ Price of the underlying,} \\
t &: \text{ Time to option's expiry.}
\end{aligned}
$$

Thus, assuming that an asset price is risk-neutral, European call option can be evaluated as:

$$C(S,t) = S\Phi(d_1) - Ke^{-rt}\Phi(d_2) \tag{1.2}$$

where

$$
\begin{aligned}
&d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}, \\
&d_2 = d_1 - \sigma\sqrt{t}, \\
&K : \text{ Option strike price,} \\
&\Phi : \text{ Standard normal cumulative distribution function.}
\end{aligned}
$$

By contrast, American option can be exercised at any time up to the maturity date, causing the asset price to be dependent of time, for which equation (1.2) becomes a free boundary problem. Given that investors have the freedom to exercise the American option at any time before its expiry, their pricing arises an optimal stopping problem to find the optimal exercise rule, then compute the expected discounted payoff under this rule. That is, at any given exercise time, the holder of an American option compares the immediate exercise payoff and the future expected payoff from continuation. Most securities traded on an exchange today are American options. These contracts will specify variables such as option premium, which is the price paid when an option is purchased or sold, as well as the strike price and the maturity date of the option.

Both European and American put and call options that have no other additional or special feature are known as vanilla, plain vanilla, normal or traditional options. In contrast, exotic options have features and conditions making them more complex. This conditions are often customized to meet a specific customer's needs. In this research we will study an example of exotic options, which is an auto-callable option.

Unlike a vanilla option contract, which remains available until its expiry date, an auto-callable is terminated automatically as soon as the underlying asset breaches a certain barrier price determined by the issuer of the note. This barrier condition is considered only for a specific (and periodic) dates, called the KO (or Call) dates. More precisely, the idea of this type of financial instrument is as follows: At specific points in time (the observation dates), it is checked whether the underlying (respectively a function of several underlyings) reaches a certain barrier. If this is the case, the buyer of the auto-callable gets a pre-defined constant cash-flow (payoff) and the certificate terminates. Otherwise the instrument continues to exist until the next observation date, and so forth. For those cases, where the auto-callable survives until maturity, a payoff depending on the underlying(s) is generated at the maturity date. All of these conditions are set based on the customers' requirements and needs.

The common point between vanilla and auto-callable payoffs is the fact that they can be exercised earlier (American Exercise type), which poses a pricing problem as no closed-form expression exists to describe the value of these options.

However, several simulation based methods have been introduced to price American options, in particular the method introduced by Longstaff and Schwartz (2001), which is easy to understand and implement as it only uses the simple method of least square. This method implements the dynamic programming principle and discrete time problem. First, they proposed the replacement of time interval of exercise dates by a finite subset. Then, the least square regression was implemented on a finite set of functions to estimate the conditional expectation function for different paths and comparing it with the payoff value of immediate exercise. Finally, starting from the last value and going backwards averaging the outcome of all paths to deduct the present value of the option, this technique is known as the least square Monte Carlo (LSM) approach.

However, the Longstaff–Schwartz method has its own shortcomings. Clément et al. (2002) proved that the the approximation function converges to the solution of the initial optimal stopping problem when the number of basis functions goes to infinity. They also proved that the normalized estimation error is asymptotically Gaussian, that is the standard error have an approximately normal distribution with mean 0 and standard deviation of $\frac{\sigma}{\sqrt{n}}$, as n the sample size goes to infinity. Moreno and Navas (2003) showed that in case of simple American option, finding basis functions used in this method is reduced to selection of degree of a polynomial on the stock price. However, the least squares regressions can face numerical problems for high polynomial degrees (more than 20). In the case of complex options, the robustness becomes questionable and the number of basis functions is not clear and its variation can affect option prices. Since the robustness and stability of the pricing is a crucial elements in risk management strategies, namely Delta hedging, we want to test if new methods can provide better pricing to the auto-callable options.

Therefore, to extend the benefits of classical methods while limiting their imperfections, a new approach is required. Reinforcement Learning (RL), in particular, is considered as an essential key access to this question, due to the ability of approximating nonlinear functions, and the aptness of simulating any policy on real data (Lagoudakis and Parr (2002); Li et al. (2009)). In fact, two reinforcement learning techniques are proposed to be promising solutions to solve this problem: the least squares policy iteration (LSPI) and fitted Q iteration (FQI) as they, in general, can easily find excellent policies within few iterations using simple sets of basis functions and samples collected from a small amount of real data (Lagoudakis and Parr (2002); Li et al. (2009)).

In this study, we want to see if reinforcement learning methods, mainly least squares policy iteration (LSPI) and fitted Q iteration (FQI), can provide better pricing of auto-callable structured products than the classical Longstaff-Schwartz method (LSM).

First, we will generate a variety of random underlying price values based on the Black-Scholes and Heston models, which is also known as Monte Carlo simulation paths. Then we apply LSPI and FQI to price auto-callable structured products and back-test the profit and lost (P&L) provided by least squares policy iteration (LSPI) and fitted Q iteration (FQI) using the historical data of Apple incorporated stock (AAPL). Finally, we will study the effectiveness of LSPI and FQI in providing better pricing of auto-callable structured products when compared to the classical Longstaff-Schwartz method.

Chapter 2

Literature review

In this chapter, we will discuss about essential elements of pricing options. First, we will briefly introduce the notion of Brownian motions, which is a pivotal concept to simulate financial dynamics. Next, we will see different models of representing value variations of underlying assets, namely Black-Scholes and stochastic volatility models. Finally, we will delve deeper into how American options can be priced using methods like the popular least square Monte Carlo method.

2.1 Diffusion Models

The term "Brownian motion" was first attributed in 1827 to the random movement of pollen grains moving in water, observed by the Scottish botanist Robert Brown (1827). This caused a major problem of mathematically expressing the movement of the pollen particles. However, it was Albert Einstein (1905) who could explain this phenomenon and solve this mystery. He did so by interpreting these irregular movement not by the classical notion of velocity, but as a stochastic process that can only be described using statistical methods. Einstein's theory of Brownian motion completely changed the understanding of these movement, and it was crucial in developing modern statistical physics as well as theories of stochastic processes. The research of finding the mathematical expression of Brownian trajectories led to the introduction of the class of stochastic differential equations as result of the work of Lemons and Langevin (2002).

On the other hand, Louis Bachelier (1900) published his thesis only a few years after Einstein's revolutionary paper. He developed the mathematical aspect of the Brownian motion opening the door to use this concept in finance and inspiring Merton (1989), which gave birth to Brownian motion models in the financial market. This later becomes the backbone of modern option pricing models such as Black and Scholes (1973) and Heston (1993) models. The definition of Brownian motion below is given by Hirsa and Neftci (2013):

Definition 2.1.1 (Brownian motion)
A Brownian motion is a random process $W(t)$, $t \in [0, T]$ such as

(a) $W(0) = 0$,

(b) $W(t)$ *has independent movements,*

(c) $W(t)$ *is continuous in* t *and*

(d) $(W(t) - W(s)) \sim N(0, |t - s|)$.

This concept will be used later in modeling the dynamics of the market and simulate price paths. The first step to understand financial situations is by representing them by abstract mathematical models, which will allow the employment of the mechanics and methods of pricing options. Many financial models are based on diffusion processes that are represented by stochastic differential equations driven by Brownian motions. In this chapter, we will be describing three models, which are Black-Scholes, local volatility and stochastic volatility models.

One of the most used models is that of Black and Scholes (1973) due to its simplicity and comprehensiveness. This model estimates the variation of assets over time and the volatility σ assumed to be constant for the entire life of the option, it is known as implied volatility.

In the Black-Scholes model, the price of the underlying asset $S(t)$ is given by:

$$\frac{dS}{S} = rdt + \sigma dW(t) \tag{2.1}$$

where $S > 0$ and $t \in [0, T)$, σ is the volatility of the stock, T is expiry date, r is the risk-free interest rate and $W(t)$ is a Brownian motion.

However, the assumption of the volatility σ being constant in the Black-Scholes model is shown to be inaccurate. Derman and Kani (1994) showed that volatility σ depends on strike price and time, their empirical results show that volatility decreases with strike level of options and increases with time to expiry date. This phenomenon is known as "volatility smile" or "smile effect".

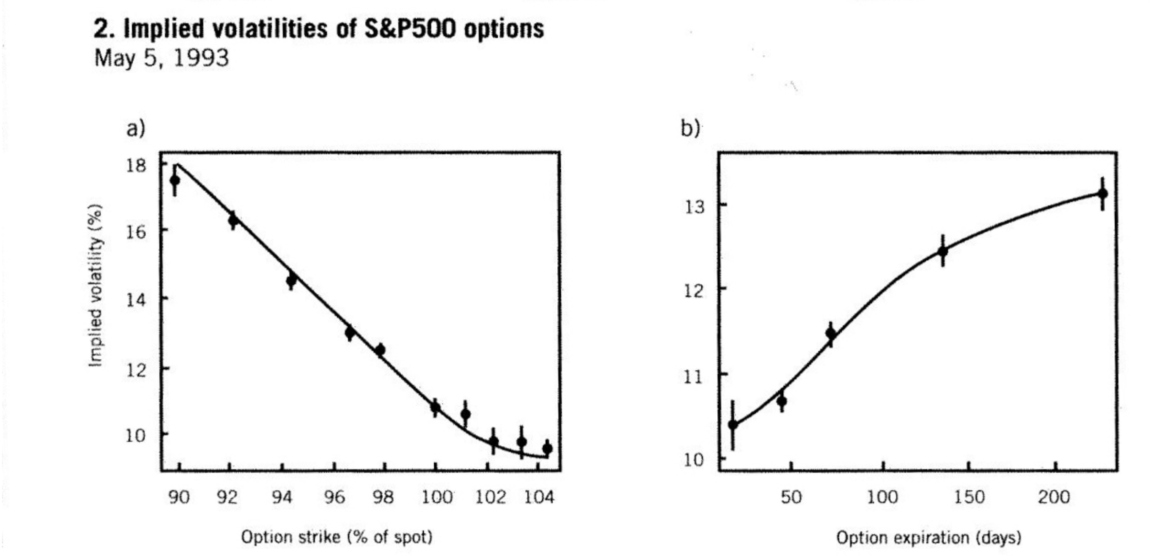


Figure 2.1: Volatility Smile by Derman and Kani (1994)

To solve this problem, Derman and Kani (1994) and Dupire (1994) incorporated volatility of the stock in the Black-Scholes model as a deterministic function of time and the price, of which this function is knows as the local volatility function $\sigma(S, t)$.

In the local volatility model, the underlying price S is given by the following stochastic differential equation:

$$\frac{dS}{S} = rdt + \sigma(S, t)dW(t). \tag{2.2}$$

Nevertheless, Hagan et al. (2002) confirmed that these models are not consistent with the empirical results, since they predict the exact opposite of the real behavior of the market. That is, local volatility models predict an increasing in the price when it actually decreases and vice versa. This problem affects the sensitivity of the option price related to its volatility, which makes the hedging processes perform worse than in the Block-Scholes model. This proves that volatility inherent some randomness and it cannot be deterministic. Different models and modification have been introduced to implement the stochastic nature of volatility (Hagan et al. (2002); Tian et al. (2015)), which are known as stochastic volatility models. One of the most popular and successful stochastic volatility models, which is not based on Black-Scholes model, is the Heston model introduced by Heston (1993). In this model, the underlying price $S(t)$ and its volatility $V(t)$ fulfill the following system of differential equations:

$$\begin{cases} dS(t) = rS(t)dt + \sqrt{V(t)}S(t)dW_1(t) \\ dV(t) = k(\theta - V(t))dt + \sigma\sqrt{V(t)}dW_2(t) \end{cases} \tag{2.3}$$

where

$\sqrt{V(t)}$: Volatility of the asset price,

θ : The long variance, that is when t tends to infinity, the expected value of $V(t)$ tends to θ,

κ : The rate at which $V(t)$ reverts to θ,

σ : Volatility of $\sqrt{V(t)}$,

W_1, W_2 : Brownian motions with correlation ρ,

q : The dividend yield.

Heston model is widely used due to its ability to capture the main features of implied volatility skews, as well as being simple to calibrate unlike other stochastic volatility models (Feng et al. (2010)). Moreover, Heston model features most of the characteristics of exotic markets, thus it is a popular choice in equity, foreign exchange and interest rate (Forde et al. (2010)), where it can be used to calculate the expectation value of a certain payoff using Monte Carlo simulations. However, it does require more parameters to be calibrated (notably κ, $V(0)$, θ, σ and ρ). Swishchuk and Vadori (2012) modified the Heston model to reduce the calibration error of its implied volatility, where the average of the absolute differences between market and the model's volatility has shown a 44% error reduction. This adjustment is made by taking into account not only the current state of the volatility, but also its previous values, which results in the delayed Heston model.

Also, Grzelak and Oosterlee (2011) discussed the case of pricing interest rate sensitive products, where the interest rate r is highly volatile. In this extension of Heston model, an additional stochastic quantity $r(S,t)$ is implemented to describe the stochastic interest rate and replace the Heston model's constant parameter r. This method also adds additional function and parameters to the original method, which may require additional calculations and more calibration errors. That is, in addition to the original two stochastic quantities in Heston model, the underlying price and its volatility, $r(S,t)$ is introduced to describe the behavior of the interest rate at every instant t. In this study we will use the Black and Scholes (1973) and Heston (1993) models, respectively.

The models, described above, are used to simulate the dynamics of the market, and will be used to create different possible paths of underlying prices. To price the option, on the other hand, different pricing methods are used, which will be discussed in the following section.

2.2 Pricing methods

For American options, the option can be exercised at any time until the expiry date, which makes the pricing process to be an optimal stochastic control problem. This is a difficult problem to solve because it does not follow the ordinary rules of calculus. For this reason many simulation based methods have been introduced to price options, notably Monte Carlo simulations (Tilley (1993); Hull and White (1987); Boyle et al. (1997); Carriere (1996)). These methods aim to calculate the possible future results based on a large number of random trials applied to a specific model. Longstaff and Schwartz (2001) used Monte Carlo simulation to estimate option values, which uses polynomial approximation of the continuation value to decide whether or not to exercise the option. Their approach, known as Longstaff-Schwartz or least squares Monte Carlo method (LSM), is one of the standard methods used for American option pricing due to its simple properties and flexibility.

However, Areal et al. (2008) proved that in case of vanilla put options, it is possible to improve the accuracy of LSM up to four times. Their method proven to be faster and more accurate due to use of a faster regression algorithm without comprising any significant loss of accuracy. Gamba (2003) used composition methods to extend LSM to value complex, compound and mutually exclusive options. Egloff (2005); Egloff et al. (2007) used non-parametric regression to estimate the continuation value and linear vector space to extend the LSM, while Kohler et al. (2010) employed neural networks to price high-dimensional American options, first generating artificial sample paths using Monte Carlo simulation, then applying least squares neural networks regression to approximate the continuation values.

On the other hand, new techniques and methods such as reinforcement learning (RL) are being introduced to solve financial problems. Reinforcement Learning is learning to take specific actions in order to maximize a numerical reward. That is, in contrast of Machine Learning where the learner is programmed to take pre-defined actions, it has actually to discover the actions that will maximize the rewards by trying different approaches until discovering the most profitable policy to take (Sutton and Barto (2018)).

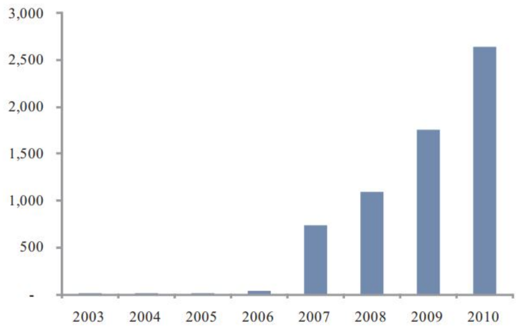
In fact, reinforcement learning techniques are gaining popularity in option pricing. For instance, Bradtke and Barto (1996) introduced a new algorithm they called least squares temporal difference (LSTD) and its recursive version (RLSTD). They showed that this method extract more information from learned experiences and converges more rapidly compared to other different learning methods. This algorithm does not have fixed control parameters, and at every iteration a better solution is discovered, thus eliminating errors caused by bad choices of parameters. Lagoudakis and Parr (2002) presented a new model-free method called least squares policy iteration (LSPI). This new method combines least squares function approximation with policy iteration which can use data gathered from any source. Another way to improve regression techniques is to use Q-learning methods which are "a simple way for agents to learn how to act optimally in controlled Markovian domains" (Watkins and Dayan (1992)). In particular the fitted Q iteration algorithm (FQI) which is capable of generalizing the capabilities of any regression algorithm, by fitting any parametric or approximation policy to the Q-function and attractively extending the horizon of optimization (Ernst et al. (2005)). In fact, it is shown that, in case of vanilla put or call options, reinforcement learning techniques can achieve better results compared to least square Method (Li et al. (2009)).

In this research, we want to examine how LSPI and FQI perform compared to LSM in pricing exotic options, especially auto-callables, as they are rapidly gaining ground in finance (Deng et al. (2016)).

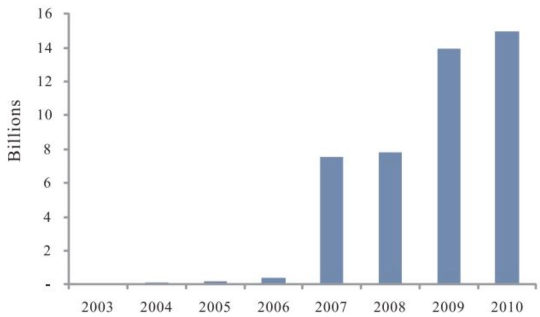
While previous studies have shown the potential of reinforcement learning in pricing traditional (vanilla put and call) American options, we want to examine their performance in pricing modern and sophisticated (exotic) type of options. In particular we will study the case of auto-callable options.

An auto-callable, which is the abbreviation of "automatically callable", is a feature of an exotic option. This feature is often found in structured products with longer maturities. A product with an auto-callable feature would be called prior to maturity by the issuer if the reference asset is at or above its initial level (or any other predetermined level) on a specified observation date. The investor would receive the principal amount of their investment plus a pre-determined premium (often paid out in the form of a coupon) and the auto-callable product is said to be redeemed early.

Exotic options are+ seen as more complex than their contemporary vanilla options. This is because they can possess specific conditions which automatically matures the option. Auto-callable products is a popular type of exotic options. That is, the option can be called automatically if the underlying passes a specific barrier on a specified observation date. Then, the investors receive the principal amount of their investment plus a coupon. The graph below represents the number and total issue size of auto-callable structured products from January 2003 to June 2010, showing a dramatic increase in the number and size of auto-callable structured products. Since this type of options is becoming increasingly common in recent years, the pricing of auto-callable options is becoming more crucial than ever before (Deng et al. (2016); Fries and Joshi (2008)).



(a) Number of Autocallables Issued



(b) Aggregate Autocallable Issue Size

Number and total issue size of auto-callables by Deng et al. (2016)

Since auto-callable are called automatically as soon as a certain barrier condition is fulfilled, they play a vital role in the volatility markets, Zvan et al. (2000) prices barrier options, where the payoff depends on the position of the underlying asset related to a specific barrier, by introducing an implicit method for solving PDE models where closed-form typically do not exist.

Deng et al. (2016) used Black-Scholes framework to model auto-callables, which yields the following partial differential equation of the option value:

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - (r + C\bar{D}S)V = 0$$

where

S : The reference asset's price,

$V(S,t)$: Price of the structured product,

r : Risk-free interest rate,

q : The dividend yield,

σ : Volatility of the underlying,

CDS : Credit default swap, which is a derivative contract which enables its owner to swap credit risk on a issuer with a counter party (the seller of the contract),

$C\bar{D}S$: Credit default swap spread, which is the annual amount the buyer must pay the seller over the length of the swap contract.

Next, this PDE is solved for both continuous and discrete call dates. For continuous auto-callables a closed-form solution has found, while the finite difference method has been used to obtain a solution with discrete call dates where no closed-form solutions exist.

Glasserman and Staum (2001), however, argues that direct Monte Carlo methods can be improved by reducing the variance of prices, that is simulated paths either cross a barrier resulting in a zero payoff or never crossing the barrier and their positive payoff can be determined. Because if the probability of not crossing the barrier is low, more simulated results in payoffs of zero. Since the average payoff is calculated only from non zero payoffs, the variance among all paths becomes relatively large. To do so, conditional distribution are employed by estimating prices conditioning on the possibility of crossing a barrier ("one-step survival").

Fries and Joshi (2008) generalises this method to a variety of product and model classes, notably non-linear boundary. They constructed Monte Carlo paths sampling only "the survival domain of the auto-callable product". Then, the pricing is calculated using finite differences of the reformulated payoffs. They show that, compared to direct simulation, this method of conditional sampling results in a reduced variance. Also, the stability of sensitivities are calculated proving that it is a more stable and consistent method. Below are the numerical results of prices and standard deviation of a Monte Carlo pricing, comparing direct simulation with conditional analytic simulation (Table 2.1), both using 5000 paths.

Product	Direct Simulation		Conditional Analytic	
Digital Caplet / Maturity t=0.5	21.40%	±0.31%	21.40%	±0.00%
Digital Caplet / Maturity t=2.0	17.38%	±0.27%	17.39%	±0.19%
Digital Caplet / Maturity t=5.0	12.04%	±0.19%	12.03%	±0.15%
LIBOR TaRN Swap 1 / Maturity t=6.0	3.56%	±0.07%	3.56%	±0.06%
LIBOR TaRN Swap 2 / Maturity t=6.05	2.511%	±0.012%	2.511%	±0.005%

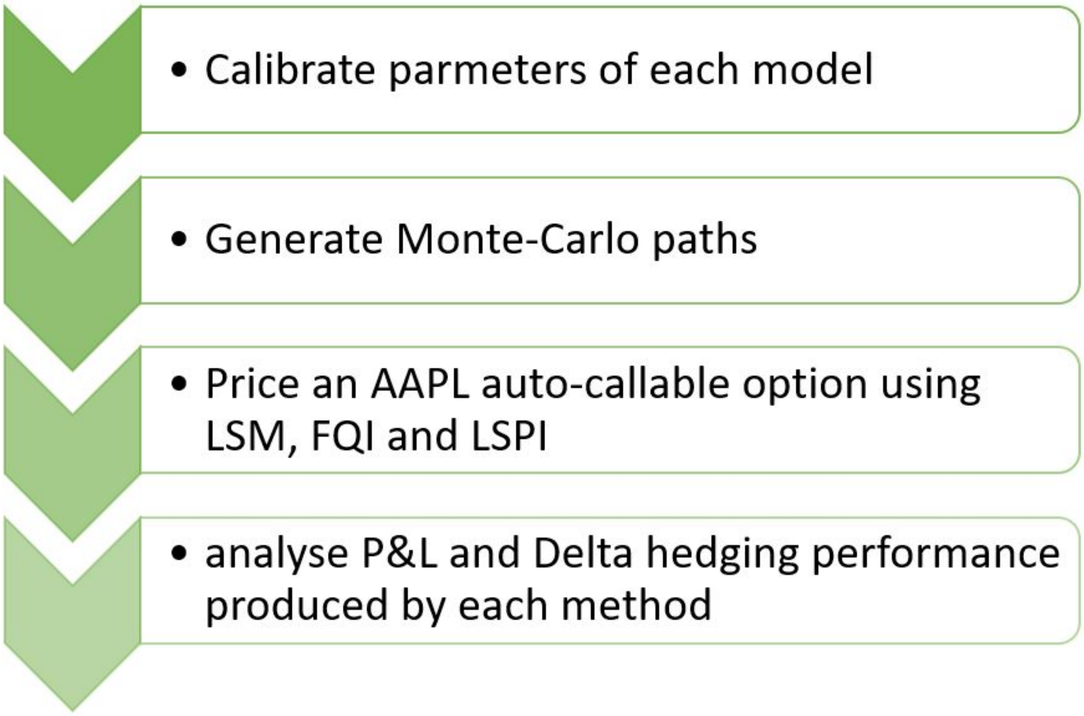
Table 2.1: Prices and standard deviation of a Monte Carlo pricing using direct simulation and conditional analytic simulation by Fries and Joshi (2008).

In this study, we will compare the Profit and Loss (P&L) produced using the standard Longstaff-Schwartz method compared to two reinforcement learning methods: least squares policy iteration (LSPI) and fitted Q iteration (FQI) in the case of an auto-callable structured product. To do so, we will use Black-Scholes and Heston models to simulate the market dynamics and we will implement Monte Carlo simulation to generate prices paths. These methods will first be tested to price American vanilla options, then we will use real data of AAPL stock (Apple Inc.) to calculate the P&L.

Chapter 3

Methodology

In this chapter, we will provide a brief introduction about different notations that we will be using for the entirety of the chapter. Next, we will elaborate more on Monte Carlo framework, that allows us to create simulated paths of underlings' values. These paths will be used later to calculate the continuation value by Longstaff-Schwartz method (LSM), also by reinforcement methods to learn the optimal trading policy. That is, using the different models we will generate a large number of different paths of underlying prices, these paths will be used by the methods to calculate the option price, this type of processes is also known as the Monte Carlo framework. The difference between the methods used in this study is characterized by how each method uses the generated paths. Finally, we will discuss the criteria of comparing the payoffs of these methods. The following flowchart illustrates the steps we will follow in order to answer the thesis question.



The methodology steps that will be taken in this study.

First, we consider a financial instrument (product) with an American exercise : The buyer of the instrument has the right to exercise the option at predefined times: $T_0, T_1, T_2, ..., T_N := T$, for a given integer $N \geq 1$. In case of exercise at the instant T_i, the buyer will receive a payoff $f(S_{T_i})$, with S_t is the price of the underlying (a stock, an index, a basket of stocks or any other fixed income or commodity underlying) and f is a function we use the payoff function $f : \mathbf{R}^d \rightarrow \mathbf{R}$, with d is the dimension of S. We assume that interest rate r is deterministic and has a flat value with respect to time.

3.1 Monte Carlo Framework

In order to simulate the market features, we will use Monte Carlo simulation to create several paths of the underlying, which then will be used to price its payoff. Monte Carlo framework is a method

for evaluating the mathematical expectation of a random variable, this method involves generating many independent samples of the random variable and then taking the empirical average of the sample as a point estimate of the expectation (Metropolis and Ulam (1949)).
The accuracy of this method is proportional to $\sigma/\sqrt{n}$ where σ^2 denotes the variance of each sample, and n denotes the number of samples generated. The key advantage of the Monte Carlo methods is that given the value of σ, the computational effort (roughly proportional to the number of samples) needed to achieve desired accuracy is independent of the dimension of the problem, i.e. if one thinks of the expectation as an integral, then this is independent of the dimension of the space where the integrand is defined. In this respect, it differs from other numerical techniques whose performance typically deteriorates as this dimension increases.

3.1.1 Simulating underlying paths

The first step in Monte Carlo framework, is the construction of a path of the underlying : an array of discrete simulation of the price S for a predefined time steps $0, dt, 2dt, ..., M \times dt := T$, with dt is a small increment of time, and M is the number of required time steps to reach the expiry T (i.e. $M := \frac{T}{dt}$. In other words, we will discretize the continuous interval $[0, T]$ into discrete M instants in order to generate different possible values of the underlying price at each instant from $t = 0$ to $t = T$. To do so, we use Euler method:

Euler Method

In this study we will use the Euler–Maruyama method as defined by Marco (2001). That is, the approximation of a continuous stochastic process X satisfying: $dX = r(t, X_t) + \sigma(t, X_t)dW_t$ is the iteration:

$$X_{i+1} = X_i + r(t_i, X_i)dt + \sigma(t_i, X_i)dWt$$

This method generates a discrete sequence $X_0, X_1, \ldots, X_M$, which approximates the process X_t on an interval $[0, T]$. The method extends the popular Euler method for numerically solving deterministic differential equations. Partition the interval $[0, T]$ into M equally spaced points.
We will use this method to create a discrete interpretation of the underlying price value S_t at each

instant $t = i \times dt$ for $i \in [|1, M|]$ based on the Black-Scholes and Heston models.

Black-Scholes Model

In the Black-Scholes Model, we use the equation:

$$dS = S(rdt + \sigma dW(t))$$

We use the fact that $W(t)$ is a standard Brownian motion as in $Definition$ 2.1.1, thus:

$$W(t + dt) - W(t) \sim N(0, t)$$
$$dW(t) = \sqrt{t}Gdt$$

with G is a random number sampled from a standard normal distribution and by applying Itô's lemma and Euler Method (Marco (2001)) we get:

$$S_{(i+1)dt} = S_{i\times dt} \exp\left\{ \left(r - \frac{1}{2}\sigma^2 \right) dt + \sigma\sqrt{dt}G \right\}$$

for $i \in [|0, M - 1|]$

Thus, generating different values of G results in creating values of S at each instant of the interval $[0, T]$, which creates different paths of the underlying price.

Heston LVSV Model

For the Heston model, we recall the system 2.3 and using Euler method we get:

$$\begin{cases} S_{(i+1)\times dt} = S_{i\times dt} \times \left(1 + r \times dt + \sqrt{V_{i\times dt}} \times \sqrt{dt} \times G_1 \right) \\ V_{(i+1)\times dt} = V_{i\times dt} + \kappa \times (\theta - V_{i\times dt}) \times dt + \sigma \times \sqrt{V_{i\times dt}} \times \sqrt{dt} \times G_2 \end{cases}$$

with G_1 and G_2 are random numbers sampled from a standard normal distribution with a correlation ρ and $i \in [|0, M - 1|]$.

Using these two models we will simulate the underlying price in order to create different paths based on the current price S_0. As seen in this both models, these paths are created with a stochastic part in form of $\sqrt{dt} \times G$. Thus, by generating random numbers sampled from standard normal distributions different paths will be constructed. That is, for a path number j, at a time step $i \times dt$ the simulated price is denoted as S_i^j. For instance, the Figure 3.1 represents 100 different price paths of an underlying using the Black-Scholes model, with $S_0 = 100$, $T = 1$, $N = 100$ and $r = 1.5\%$

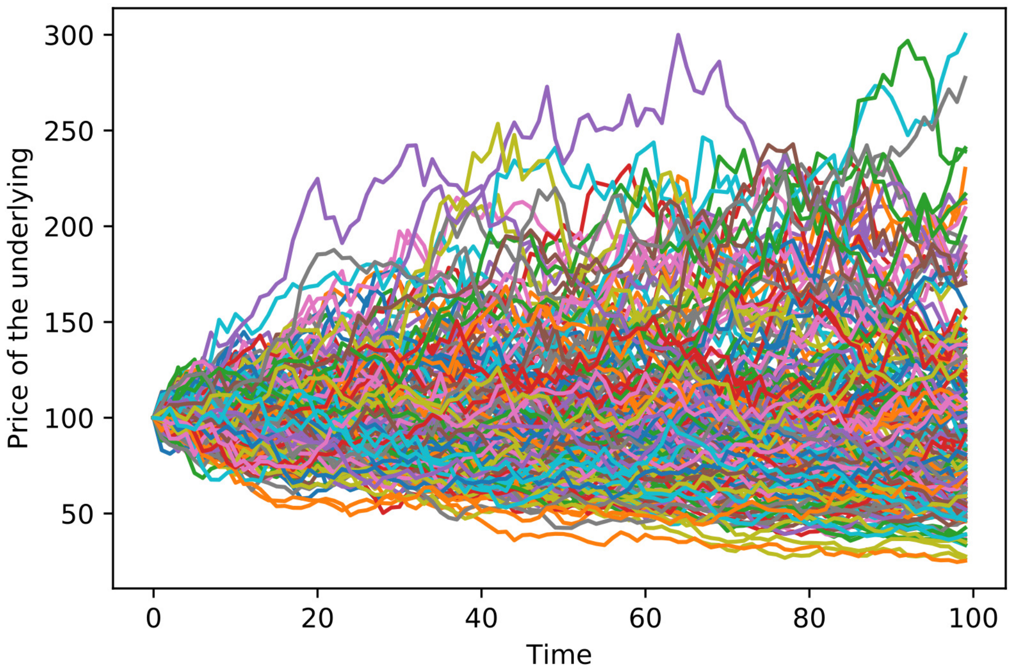


Figure 3.1: Monte Carlo Grid representing 100 different path realizations from the original value $S_0 = 100$ showing possible paths of the underlying price.

These paths will be used in the following methods to calculate the continuation value at each instant, which is the expected payoff if the investor decided to hold the option rather than executing it.

3.2 Continuation Value

In order to find the optimal execution rule which defines the exercise date T, the buyer of the instrument constantly compares the current value to the expected future value of the option. In other words, to approximate the stopping rule we need to estimate the conditional expectation $E[f(S_{t+dt})|S_t]$, also known as the continuation value, of the option value at time $t + dt$ conditional on its value on the previous instant t.

3.2.1 Longstaff-Schwartz

The Longstaff-Schwartz method aims to estimate the conditional expectation $E[f(S_{t+dt})|S_t]$ at each instant t with a set of basis functions using least squares regression:

$$\widehat{E}[f(S_{t+dt})|S_t] = \sum_{i=0}^{N} a_i L_i(S_t)$$

where these functions can be chosen as a set of polynomials, Longstaff and Schwartz (2001) proposed the use of weighted Laguerre polynomials, defined by:

$$\begin{aligned}
L_0(X) &= e^{-X/2}, \\
L_1(X) &= e^{-X/2} \quad (1 - X), \\
L_2(X) &= e^{-X/2} \quad (1 - 2X + X^2/2), \\
L_n(X) &= e^{-X/2} \quad \frac{e^X}{n!} \quad \frac{\mathrm{d}^n}{\mathrm{d}X^n}\big(X^n e^{-X}\big).
\end{aligned}$$

This procedure runs backwards in time for every path starting from $t = T, T-1, ...t = 1$, when we regress the value of the payoff $f(S_t)$ against the N first basis functions $L_n(S_t)$. At the end of this part, we will have an estimation of $\widehat{E}[f(S_{t+dt})|S_t]$. Then we start the grid from the beginning at $t = 1$ and move forward in time to find the optimal trading strategy. That is, finding the execution time T that maximises the payoff of the option $f(S_T)$. To do so, at each instant, we compare the value of the immediate payoff $f(S_t)$ with $\widehat{E}[f(S_{t+dt})|S_t]$. The option should be executed as soon as $f(S_t) >= \widehat{E}[f(S_t)|S_{t-dt}]$, and then we found the execution time $t - T$.

3.2.2 Reinforcement learning algorithms

While the previous method is widely used to price American options, some alternative methods are promising to be more efficient and accurate, especially the ones based on reinforcement learning algorithms. To introduce these methods, we will have to define a policy π and an action-value function Q:

A policy π is a rule to for taking actions based on stated visited. $\pi(s,a)$ is the probability of choosing an action a in a state s while following the policy π. An optimal policy maximizes the expected total discounted rewards obtained, so the action-value function Q is given by:

$$Q^{\pi}(s,a) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t | s_0 = s; a_0 = a\right]$$

where $\gamma \in [0,1)$ is the discount factor and r is the reward received when transitioning from state s to next state s' taking the action a. The optimal action-value function satisfies:

$$Q^*(s,a) = \sup_{\pi} Q^{\pi}(s,a).$$

Since options can be exercised one time only, the reward r and the action a will have only two possible values, namely 0 and 1. We will put $r = 0$ and $a = 0$ until the execution date where $a = 1$ and $r = \gamma f(S_t)$. The state s denotes (S_i, i) , where S_i the underlying price and the time step i.

We can approximate the function Q with a linear combination of some basis functions $\phi = (\phi_1, ..., \phi_d)^{\intercal}$ and its weight $\omega = (\omega_1, ..., \omega_d)^{\intercal}$, where : $Q(S_i, 0) = \omega^{\intercal}\phi(S_i, i)$ and $Q(S_i, 1) = f(S_i)$. In other words, the option will be executed only if the immediate payoff is greater than the continuation value. First, we will estimate the continuation value with $\omega^{\intercal}\phi(S_i, i)$ and at each iteration i, we will compare it with the immediate payoff $f(S_i)$. The option will be exercised as soon as $\omega^{\intercal}\phi(S_i, i) < f(S_i)$. The policy π is, therefore, defined as $\pi(S_i) = 1$ if and only if $\omega^{\intercal}\phi(S_i, i) < f(S_i)$.

Since different methods take different approaches to calculate the weight ω, we will discuss each method separately.

Least squares policy iteration

The first method is the least squares policy iteration, which is a method of discovering the optimal policy for any given Markov Decision Processes (MDP). It is an iterative procedure in the space of deterministic policies to discover the optimal policy by generating a sequence of monotonically improving policies. LSPI algorithm arbitrary chooses initial arguments, then for each path j (with $j \in [|1, N|]$) we compute the state-action value function $Q^{(j)}(s,a) = \left(\omega^{(j)}\right)^\intercal \phi(s,a)$. By solving an equation of the the following form:

$$A^{(j-1)}\omega^{(j)} = b^{(j-1)}$$

A and b are two method parameters and the initial values of w, A and b are set to be zero. In each iteration if $\pi(S_i) = 0$ the value of A is updated by $\phi(S_i^{j-1}, i)\left(\phi(S_i^{j-1}, i) - \gamma\phi(S_{i+1}^{j-1}, i+1)\right)^\intercal$, and $b = 0$. On the other hand, if $\pi(S_i) = 1$ we update the values of A and b by $A = \phi(S_i^{j-1}, i)\left(\phi(S_i^{j-1}, i) - \gamma 0\right)^\intercal$, and $b = \gamma f(S_{i+1}^{j-1})\phi(S_i^{j-1}, i)$ Thus, we can compute A and b with the following formula:

$$\begin{cases} A^{(j-1)} = \sum_i \phi(S_i^{j-1}, i)\left(\phi(S_i^{j-1}, i) - \gamma \mathbb{1}_{\left[\pi_{j-1}(S_{i+1}^{j-1})=0\right]}\phi(S_{i+1}^{j-1}, i+1)\right)^\intercal, \\ b^{(j-1)} = \gamma \sum_i \mathbb{1}_{\left[\pi_{j-1}(S_{i+1}^{j-1})=1\right]}\phi(S_i^{j-1}, i) f(S_{i+1}^{j-1}). \end{cases}$$

Fitted Q iteration algorithm

The second method is called the fitted Q iteration algorithm. It is a batch mode reinforcement learning algorithm which yields an approximation of the Q-function corresponding to an infinite horizon optimal control problem with discounted rewards, by iteratively extending the optimization horizon, thus:

$Q^{(j)}(s,a) = \left(\omega^{(j)}\right)^\intercal \phi(s,a)$ and $A^{(j-1)}\omega^{(j)} = b^{(j-1)}$, where

$$\begin{cases} A^{(j-1)} = \sum_i \phi(S_i^{j-1}, i)\phi(S_i^{j-1}, i)^\intercal, \\ b^{(j-1)} = \gamma \sum_i \max(f(S_{i+1}^{j-1}), \omega^\intercal \phi(S_{i+1}^{j-1}, i))\phi(S_i^{j-1}, i). \end{cases}$$

3.3 Market and Model Risks

After calculating the continuation value, we can obtain a clear trading strategy, which can give us the payoff value. After getting the structures' price, it is important to analyze the different factors that can affect its stability including risk measures known as the 'Greeks', model and calibration risks.

3.3.1 Reminder on the sensitivities: the Greeks

A structure's price can be influenced by a number of factors that can either help or hurt traders depending on the type of positions they have taken. Successful traders understand the factors that influence the product pricing, which include the so-called 'Greeks'—a set of risk measures so named after the Greek letters that denote them, which indicate how sensitive an option is to time-value decay, changes in implied volatility, and movements in the price its underlying security. Moreover, 'Greeks' play an important role in Delta hedging, which is a trading strategy that aims to reduce the directional risk related to the asset movement. The objective of this strategy is to purchase or sell certain number (equals to Delta) of shares in order to offset the total risk of the portfolio.

Delta

$\Delta := \frac{\partial P}{\partial S}$ represents the fist-order sensitivity of the derivative's price to a move in the underlying price S. For a small change move ϵ in S, the price of the structure will move by $\epsilon \times \Delta$. To hedge against movements in the underlying asset, the buyer of a call option must buy Delta units of the underlying, he would then obtain a Delta-neutral portfolio that has to be hedged dynamically in order to remain Delta-neutral throughout the option's life.

Gamma

$\Gamma := \frac{\partial^2 P}{\partial S^2}$ represents the second-order sensitivity of the option to a movement in the underlying asset's price. It gives a second-order correction to Delta, as the option is a non-linear function of the underlying price. It also underlines the first-order sensitivity of Delta to a movement in the underlying.

Vega

$Vega := \frac{\partial P}{\partial \sigma}$ represents the sensitivity of the option price to a movement in the volatility of the underlying asset. This sensitivity contradicts the essential Black-Scholes assumption on constant volatility, but is absolutely indispensable in order to manage a book of derivatives. The Vega is greatest around the strike price K and falls exponentially on both sides.

Theta

$\theta = \frac{\partial P}{\partial t}$ represents the rate at which the option price varies over time, on a daily basis. That is, the buyer of a call or a put option will have a negative Theta: as time passes, the option has less time to expiry and therefore loses time value.

Rho

$\rho := \frac{\partial P}{\partial r}$ represents the sensitivity of an option price to a movement in interest rates. As interest rates only have a first-order impact on the option price, call and Put options are almost linear in interest rates. This effect is linked to a dual influence of interest rates, firstly on the cost of Delta hedging, and secondly on price discounting.

Vomma

$Vomma := \frac{\partial^2 P}{\partial \sigma^2}$ represents the second-order sensitivity of the option price to a movement in the implied volatility of the underlying asset: it corresponds to options that are convex in volatility, that is to say options that have Vega convexity.

Vanna

$Vanna := \frac{\partial^2 P}{\partial \sigma \partial S}$ represents the sensitivity of the option price to a movement in both the underlying asset's price and its volatility. It can be seen as the sensitivity of an option's Delta to a movement in the volatility of the underlying. It is a good indicator of how much a Delta hedge is going to change if volatility moves.

In this study, however, we will primarily focus on Delta and Gamma values as they play the main roles in the Delta hedging strategy.

3.3.2 Monte Carlo Risks

Monte Carlo Framework has two main sources of error that a trader should take into account when pricing a structure, time discretization error schemes as discussed in Section 3.1.1 and the sampling error.

Discretization error

The path dependency of auto-callables requires a sufficient number of steps to accurately model price evolution. For example, the stock price simulation. In fact, Janzon (2018) proved that in order to attend a precision of ϵ, the number of steps m should satisfy $m = \mathcal{O}(\frac{1}{\epsilon^2})$. Thus, if a fewer number of steps are is used, a barrier might not be triggered which would otherwise have been triggered if more number of steps were used. That is, increasing number of steps provides a better modeling of the underlying's price movement, thus increasing accuracy of the simulation. This becomes very important while pricing path dependent derivatives.

Crude Monte Carlo

In order to get an acceptably accurate estimate of the option price, very large number of simulations has to be performed, typically in the order of millions. This problem can be dealt with using variance reduction methods. These methods work on exactly the same principle as that of hedging an option position, that is that the pay-off of a hedged portfolio will have a much smaller variability than an unhedged pay-off. This corresponds to the variance (or equivalently standard error) of a simulated hedge portfolio being much smaller than that of the unhedged pay-off.

3.3.3 Calibration Risks

Option pricing models are calibrated to market data of plain vanillas (call and put options) by minimization of an error functional. From the economic viewpoint, there are several possibilities to measure the error between the market and the model. These different specifications of the error give rise to different sets of calibrated model parameters and the resulting prices of exotic options vary significantly. These price differences often exceed the usual profit margin of exotic options.

Black-Scholes model requires the calibration of the volatility around the strike price K. At this price, the option value using the model should equal the market value, this yields a unique value of volatility σ as the option value is strictly increasing with respect to σ. For the Heston model, we have more degree of freedom, so we can calibrate more data than just an ATM option. If P is the value of the option, σ is calibrated as in Black-Scholes model to match the market price. The correlation ρ is calibrated to match the option skew, which is ΔP of the market should equal ΔP of the model and the volatility of σ can be calibrated by getting the historical standard deviation of option value, this standard deviation then should match the model volatility of σ.

In this chapter we have covered the models we will use to simulate the underlying price, namely Black-Scholes and Heston models. These methods will provide formulas to generate Monte-Carlo simulation paths, which will be used by the pricing methods to valuate options. In the next chapter, we will define and price a particular auto-callable note, then we will analyze the P&L and hedging performance of each of the methods.

Chapter 4

Auto-callable-structured option

In this chapter we will calibrate and calculate the parameters of Black-Scholes and Heston models using American vanilla put options. We will use 20,000 training paths and 10,000 for testing. Also we will use the interest rate $r = 2.5\%$ and $\sigma = 39\%$, maturity of the option is $T = 1$, and number of steps $m = 100$. Then we will consider an exotic type of options, which is auto-callable-structured option. First, we will remind of its definition and notations. Next, we will discuss the fair coupon calibration process. Finally, we will price and examine the Profit and Loss performance of LSM, LSPI and FQI using real data of Apple Inc. (AAPL Stock) from January 14, 2019 to January 13, 2021.

4.1 Black-Scholes Model

Black-Scholes model is one of the most used models to express dynamics of economical derives. That is due to its simplicity and comprehensiveness. This model represents the variation of an asset following the Equation 2.2. We will use Black-Scholes model to simulate the behaviour of the underlying, in order to price American vanilla put options and its Greeks, and it will be used later to calibrate more complicated models such as the Heston model. The following table represents the price of American put options using LSM, LSPI and FQI utilizing Black-Scholes model, where S_0 is the underlying price and the strike price is $K = 140$.

S_0	LSPI	FQI	LSM
100	42.73	42.73	42.73
110	35.77	35.77	36.05
120	29.76	29.76	29.77
130	24.84	24.84	24.84
140	20.62	20.62	20.63

Table 4.1: In the money

S_0	LSPI	FQI	LSM
140	20.62	20.62	20.63
150	16.20	16.20	16.67
160	13.35	13.35	13.35
170	10.43	10.43	10.64
180	8.61	8.64	8.62

Table 4.2: Out of the money

To test all parameters of the model and methods, we can see that all three methods give similar prices. Also we test the three methods who should always verify the non-arbitrage boundary, which states that American option value is greater or equal than its contemporary European value. This inequality should hold for every model, which can reassure that all parameters are correctly calibrated. For instance we look at the following table representing prices of European put option, having identical parameters, with respect to the spot price S_0:

S_0	100	110	120	130	140	150	160	170	180
P	41.97	34.91	29.22	23.77	19.57	16.04	13.29	10.41	8.60

Table 4.3: European put option prices

As expected, all the 3 methods above are producing higher prices of American options than their European counterparts for all spot scenarios. In contrast to European-style options which are time limited, American option holders have more freedom in exercising their options, which offers them a lot more payoff possibilities. Therefore, the price of American options is always higher than the value of European options. Under the Black-Scholes model, we use LSM, LSPI and FQI to calculate the Delta ($= \frac{\partial P}{\partial S}$) values of spot prices around the strike K:

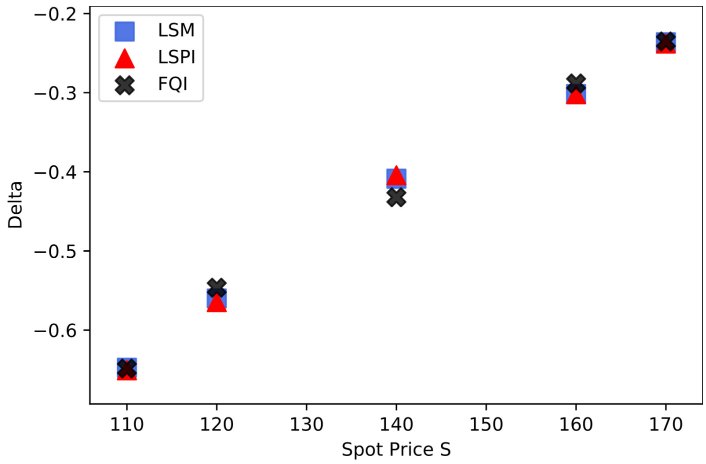


Figure 4.1: Delta values with respect to to S_0

We remark that all the 3 methods are giving close levels of Delta. The curve is increasing with respect to to S_0 as expected: When the spot price is low, the intrinsic value (which has a Delta=-1.0) has more contribution on the option price, and the put payoff is converging to a simple spot spread,i.e. $K - S_0$. Delta remains negative for all spot scenarios as the put option can just decrease when the spot moves up.

4.2 Heston Model

Heston stochastic model is used for highly volatile assets, and it will be used to price auto-callable-structured options. First we will use Black-Scholes results in order to calibrate its parameters. Since Heston and Black-Scholes model should produce the same European option prices, we can calibrate the Heston parameter V_0 using Newton–Raphson method as described by Gil et al. (2007). To do so, we fix underlying price at $S_0 = 200$ for example and the strike price as $K = 140$ (which yields $f_{BS}(V_0) = f_H(V_0) = 5.15$), and we want to find the Heston parameter V_0. If we denote the European put option price provided by Black-Scholes and Heston models as $f_{BS}(V_0)$ and $f_H(V_0)$ respectively, the calibrated V_0 represents the root of the function $f_{BS}(V_0) - f_H(V_0)$. Figure 4.2 illustrates the calibration process of V_0.

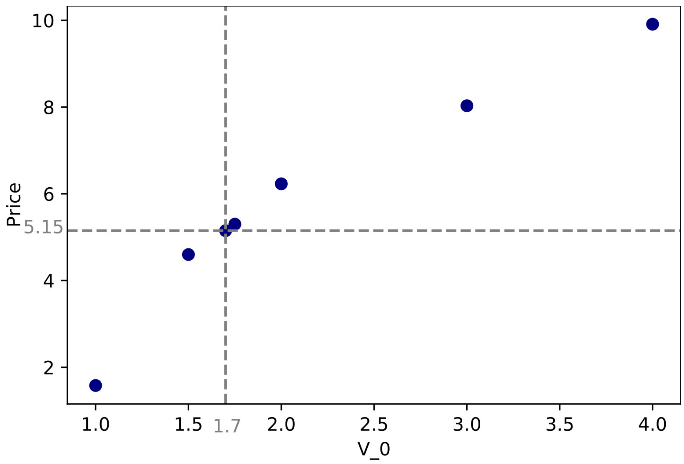


Figure 4.2: Heston model calibration where we change the value V_0 until we get the market price $P = 5.15$.

After this process we obtain $V_0 = 1.7$, then we proceed to calculate the option price at this point. The following table represents the price using the three different methods:

	LSM	LSPI	FQI
Price	5.3	5.28	5.44
Delta	0.58	0.55	0.57
Gamma	0.024	0.027	0.022

Table 4.4: Option Prices for K=140

Similar to the Black-Scholes model, we obtain relatively close option, Delta and Gamma values, which suggests that LSPI and FQI are at least as efficient as LSM in pricing vanilla options.
Next we will price auto-callable-structured options.

4.3 An Auto-Callable Note

Let us begin by describing in detail the payoff structure, that will be used in this study, without the callability option (i.e. the underlying of the structure). We consider the expiry of the payoff to be 1 year, with 4 quarterly observation ($T_1 = 3$ Month, T_2 =6 Month,$T_3 = 9$ Month and $T_4 := T = 1$ Year). The holder of the structure pay some coupon amount c at each observation $T_i,\ i \in \{1, 2, 3\}$ as :

- At T_1 if the observed stock price is above the barrier $120\% \times S_0$ the holder pays c and the structure continue.Otherwise the the holder receives the stock performance S_{T1}/S_0.
- At T_2 if the structure is still alive (from T_1) and if the observed stock price is above the barrier $120\% \times S_0$ the holder pays c and the structure continue. Otherwise the the holder receives the stock performance S_{T2}/S_0.
- At T_3 if the structure is still alive (from T_2) and if the the observed stock price is above the barrier $120\% \times S_0$ the holder pays c and the structure continue. Otherwise the the holder receives the stock performance S_{T3}/S_0.
- Finally at last expiry T if the structure is still alive, the holder receives unconditionally the performance S_T/S_0.

If we denote payoff as PO, we get:

$$PO = -c\sum_{i=1}^{3} \mathbb{1}_{A_i} + \sum_{i=1}^{3} \mathbb{1}_{\bar{A}_i} \times S_{Ti}/S_0 + \mathbb{1}_{A_T} \times S_T/S_0 \tag{4.1}$$

with $\mathbb{1}$ is the indicator function and $A_i, i \in 1, 2, 3$ are the auto-callable barrier conditions (also known as the knock out events) because as soon as one of them is fulfilled, the option is called automatically (or knocked out):

- $A_1 = \mathbb{1}_{S_{T1} < 120\% \times S_0}$.
- $A_2 = \mathbb{1}_{S_{T2} < 120\% \times S_0} \times \mathbb{1}_{A1}$
- $A_3 = \mathbb{1}_{S_{T3} < 120\% \times S_0} \times \mathbb{1}_{A2}$
- $A_T = \mathbb{1}_{\bar{A}_3}$

The next section shows how the coupon amount c is fixed.

4.4 Fair Coupon Calibration

The value of the coupon c at the initiation date will be zero to both parties. For this statement to be true, the values of the cash flow streams that the holder pay should cancel the expected performance to receive :

$$0 = E[PO] = -c\sum_{i=1}^{3} E[e^{-rT_i}1_{A_i}] + \sum_{i=1}^{3} E[e^{-rT_i}1_{\bar{A}_i} \times S_{Ti}/S_0] + E[e^{-rT}1_{A_T} \times S_T/S_0]$$

We use Newton–Raphson method to calibrate the value c, as it represents the root of $E[PO]$. The fact that the price of the underlying note is a strictly decreasing function with respect to the coupon amount ensures the existence and uniqueness of the fair coupon. The graph below shows how the price under Black-Scholes is changing with levels of c. We can see effectively that $c = 0.46$ is giving the fair level for this model.

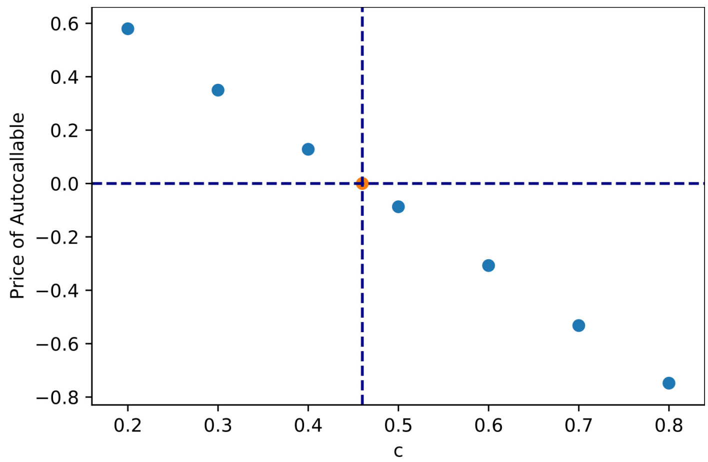


Figure 4.3: Black-Scholes Auto-callable note prices with respect to. the coupon amount c.

In the next test, we choose use the calibrated Heston model as discussed in Section 4.2. Figure 4.4 below shows the fair coupon level for this model :

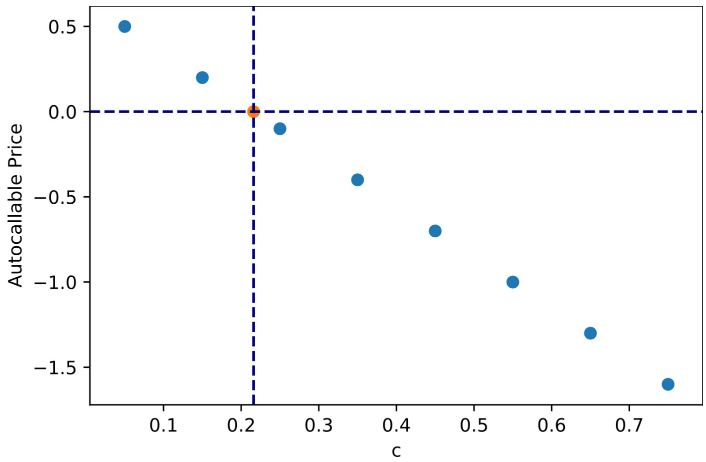


Figure 4.4: Heston Auto-callable note prices with respect to. the coupon amount c.

We can see that the fair coupon is slightly lower in Heston model than in Black-Scholes, even if the first model's volatility was calibrated to produce the same Put option price. This is an expected result because Heston model is correlating the spot with the volatility which is creating a negative skew, the barrier option included is in this product is sensitive to the move of the volatility with respect to the spot, a negative correlation is introducing more variance on the left side of the digit which is giving a higher price of the structure. Thus, the fair coupon should be reduced to compensate this higher price (relatively to Black-Scholes).

Another way to see this skew effect introduced by Heston, is to check the fair coupon level changes when the spot/volatility correlation in Heston model is changing (i.e. the parameter ρ).

Figure 4.5 shows how the coupon changes when the correlation is changing from -0.8 to 0.8. We can remark the correlation has a symmetric effect on the coupon price (coupon level less expensive for a purely de-correlated volatility) and it is taking the maximum level for highly correlated state (both 100% and -100%).

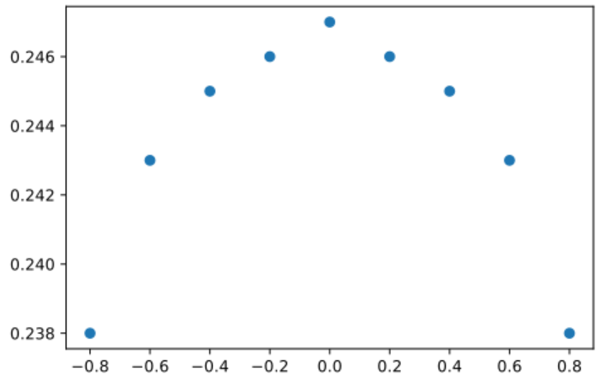


Figure 4.5: The value of the fair coupon amount c related to the Heston model with respect to ρ.

Figure 4.5 shows that the correlation ρ has a direct impact on the coupon value, which has a significant impact on the option pricing. Since the Black-Scholes model implies that the correlation $\rho = 0$, it does not reflect the impact of the correlation on the coupon price. We conclude that Heston model is more appropriate to price this structure as the market volatility has a skew effect (see Figure 2.1) and this effect is taken into account with Heston correlation parameter ρ (Hull (2003)). In other words, the coupon value c is directly impacted by the volatility of the underlying and its changes, consequently the auto-callable price is highly sensitive to the volatility.
In the following sections, we will use a pre-calibrated level of $\rho = -57\%$.

4.5 Pricing The Callable Structure

After having calibrated the coupon level, in this section we are able to price the callable option: The buyer of the product has the right to cancel the structure at any exercise time among T_1, T_2, T_3. The pricing of this structure requires the modeling of the exercise boundary in a similar way as we had seen for the American put option , in fact at each exercise date T_i the holder can decide either to cancel or continue the structure depending on the future expected value of the product, let the continuation value noted as C and the future payment is f, so $C_{T_i} := E_{T_i}[f]$. One can distinguish two scenarios :

- If $C_{T_i} > POf_{T_i}$ then the holder prefers to continue the structure.

- If not, then it is more advantageous to exercise the callablility option and stop the product.

with POf_{T_i} is the underlying auto-callable note payoff as described in the previous section. As in the American put option, the continuation value will be estimated using the 3 methods LSM, FQI and LSPI. Figure below is showing the callable option price for different scenarios of spot between 125 and 150.

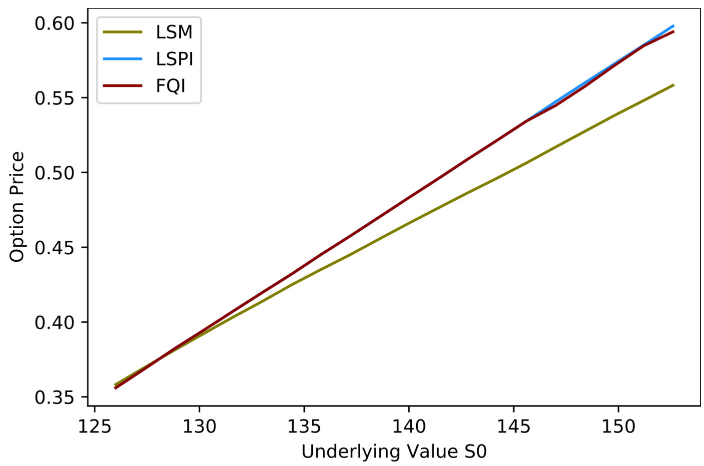


Figure 4.6: Callable structure FV with respect to spot moves for the 3 methods LSM, LSPI and FQI

We remark that LSPI and FQI are giving very close prices, while the classical LSM is slightly less expensive (especially for high values of S_0). This is explained by a common issue on LSM: The method is under-estimates the callable option, because of the call digital irregularity (i.e. discontinue Delta) which cannot be perfectly explained by a polynomial form. LSPI and FQI are, on the other hand, providing a more complex form of regression, which can estimate better the call irregularity. In order to point up this advantage of LSPI and FQI upon the classical LSM method, we will back-test the P&L (Profit and Loss) of a seller of this product within the last 2 year on AAPL stock (Apple Inc.).

4.6 BackTest and P&L performances

In this section, we will examine a realistic scenario based on the last 2 years historic data of AAPL stock. Figure 5.5 is showing the daily close prices of this stock since t_0 =14-Jan-2019

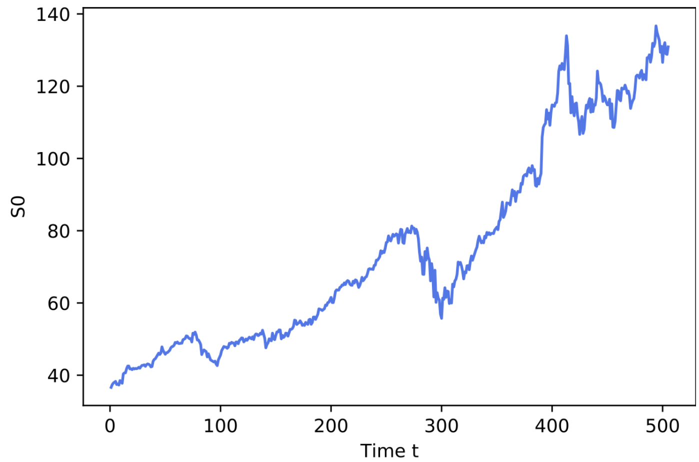


Figure 4.7: AAPL historical daily spot prices from Jan-2019 to Jan-2021.

For every day t and starting from the first day t_0, the seller of the callable option will use either LSM, LSPI or FQI to produce a price FV_t^x to the holder (with $x \in \{LSM, LSPI, FQI\}$). Each method is producing a Delta sensitivity as well Δ_t^x. The seller will use the Delta to hedge her position against any spot move, this is done by purchasing a number Δ^x of AAPL shares S_t.

The P&L (Profit and Loss) of the seller, for one day period from t to $t+1$, is constituted of the change of the option price $-FV_{t+1}^x + FV_t^x$ and the hedge position $\Delta_t^x \times (S_{t+1} - S_t)$. The total daily P&L writes then :

$$P\&L_t^x = -(FV_{t+1}^x - FV_t^x) + \Delta_t^x \times (S_{t+1} - S_t)$$

The objective our analysis here is to study the P&L distribution for each method.

The Figure below shows the daily hedged P&L for the 3 methods of AAPL stock option. We can oberve that LSPI and FQI are giving more consistently stable results. That is the hedged P&L boundaries of the two reinforcement learning are always smaller than those given by LSM.

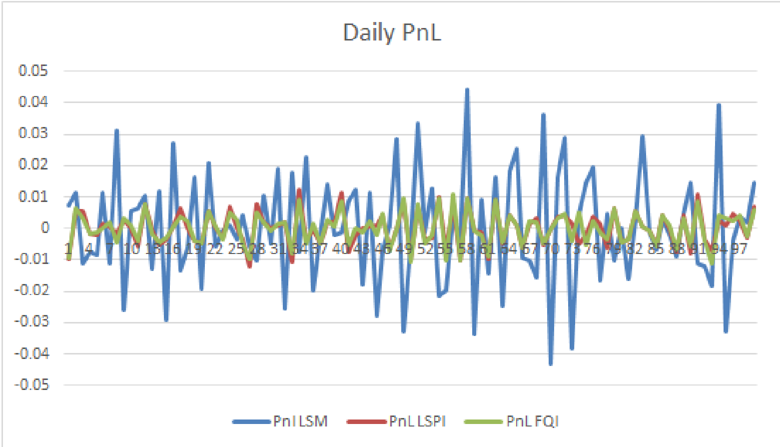


Figure 4.8: Daily hedged P&L for each method LSPI, FQI and LSM .

This stability improvement can be measured with the standard deviation, where a more stable daily P&L represents a smaller standard deviation. Table 4.5 shows a ratio of 1/3 between the new methods and the classical LSM, this is implying a less risky and more stable performance with this two methods. For LSPI and FQI, we obtain mean levels of P&L with small values (less than 10^{-5})). Thus, the pricing and hedging using the new methods is significantly more precise by means of auto-callables.

	LSM P&L	LSPI P&L	FQI P&L
Mean	1.11×10^{-4}	1.84×10^{-5}	-1.39×10^{-6}
Std Dev	1.78%	0.54%	0.53%

Table 4.5: Average and Std Deviation of daily P&Ls for each method.

Finally, we examine the histogram of the P&L for each method, i.e. the probability distribution of the variable $P\&L$ as it was realized each day (based on our back-test). Figure 4.9 shows a more concise distribution of LSPI and FQI than LSM : 30% of the realisations under LSPI and FQI gives a perfect hedge state (i.e. 0.0). While more than 75% of the $P\&L$s are between -0.005 and 0.005. On the other hand, only 35% of the outcomes within LSM method are in the same interval -0.005,+0.005, and 75% of the realisation under this method are between -0.02 and +0.02. This represents a confidence-interval 4 times wider than the one given by FQI and LSM.

From a financial perspective, this result suggests that whenever a trader hedge his product, he has a

75% chance that his portfolio moves with $+/-0.005\$$ the day after if he is using the new methods. But if he is using LSM then the 35% is the chance to be within the same hedge error of$+/-0.005\$$.

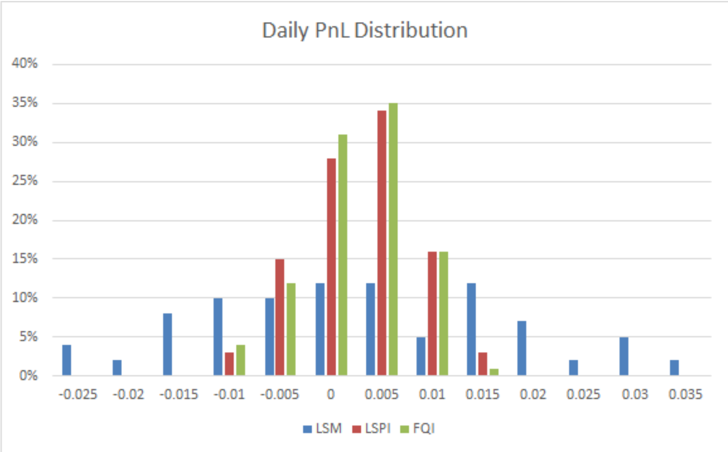


Figure 4.9: P&L distribution Histograms for LSM, LSPI and FQI.

The previous results show that, while all the three methods provide relatively similar vanilla option prices, LSM significantly underestimates the value of the callable option compared to LSPI and FQI. This can be explained by the fact that vanilla options have continuous payoff functions, which can be easily estimated using least squares regression.

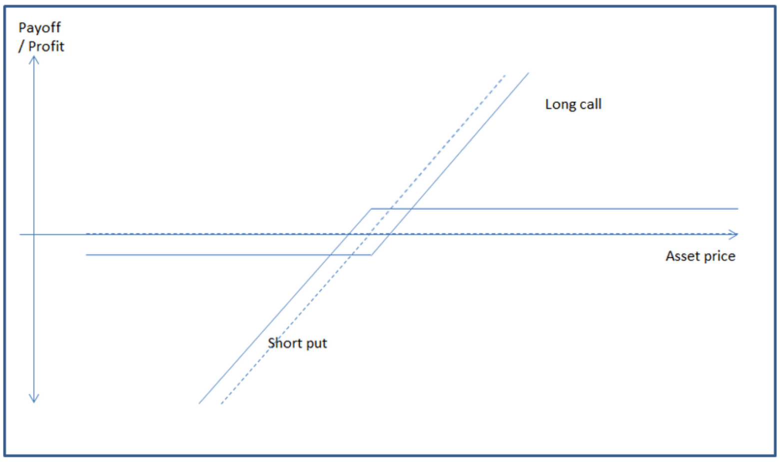


Figure 4.10: Payoff of vanilla options by Zhao (2018), showing the continuity of the payoff function.

On the other hand, structured options, such as the auto-callable used in this study, have a discontinuity of the payoff. Such discontinuity cannot perfectly projected on polynomial basis (or any other continuous function basis).

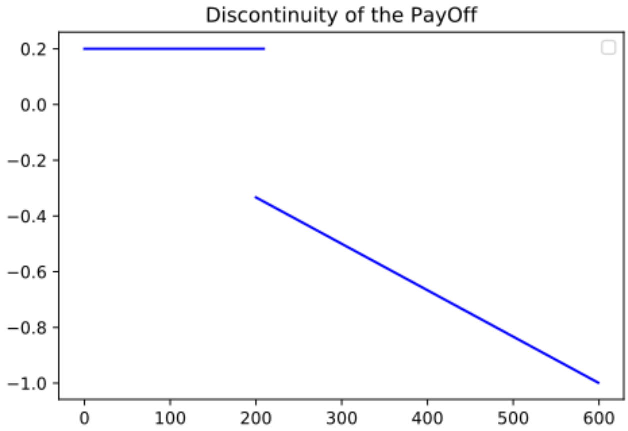


Figure 4.11: Payoff of structured options, showing a discontinuity of the payoff function.

On the other hand, the new methods LSPI and FQI are providing a better degree of complexity which is needed to explain the discontinuity of the payoff. That is, reinforcement learning methods can extract valuable information from any set of data, and the recursive process allows the improvement of the approximations after every irritation (Bradtke and Barto (1996)). Also, Hachiya and Sugiyama (2007) proved that Reinforcement learning methods can achieve very efficient results dealing with discontinuities, where they suggested the use a new basis functions called geodesic Gaussian kernels to overcome the discontinuity problem. The complexity to catch-up by the model is even higher for Delta risk (as it's a derivative of a discontinuous function and as consequence, it should represent some irregularities), this explains why the new methods are giving a more stable hedge and by consequence a smaller std deviation of the daily hedged moved, a very important results for a trader who is seeking to perfectly cancel the Delta risk of her portfolio.

Chapter 5

Conclusion And Future Work

In this project we had explored two new alternatives of the classical LSM approach to price American style options. The new approaches are based on reinforcement learning technique coupled under Monte Carlo technique.

Firstly, the comparison is applied to a simple put option, in which we show the new approaches are preserving the same price and same Delta profiles as LSM under different spot scenarios. Secondly, we extended this comparison to a more complicated structure : the callable option of an auto-call note. This study shows how Heston model is giving a more appropriate price-by considering the skew effect-than the basic Black-Scholes model. However, it shows how the new FQI and LSPI methods are ameliorating the classical underestimation of the exercise boundary by LSM.

Finally, we applied the three pricers to a realistic data and provided an analysis of the performance of a Delta hedged position on the callable option. This analysis is showing how the new methods improvement of the exercise boundary brings a better hedging performances (3 times lower standard deviation and more precise average).

In a future study, an analysis of other risks, especially Vega risk, as the instrument is highly sensitive to long and short dated volatility, which requires not just a stable total Vega but also a smooth Vega term-structure (i.e. Vega with respect to different exercises dates T_i), because a trader will try to offset the Vega risk by buying a stream of vanilla options expiring as the exercise dates. The total portfolio should have very negligible Vega risk but also the residual risk should be stable from a day to another (so the trader is not losing money to re-balance his Vega very frequently, especially that a transaction cost is applied whenever the quantity of the hedging options are adjusted).

We can also suggest another direction to extend this study : Apply the same pricing and Delta hedge back-tests to different asset type than equities, for examples Interest Rates or Credit based Derivatives, for which the volatility skew shape is different and the resulted risks are more complicated (for examples for interest rate products, the underlying is the rate forward curves , which means the related Delta risk is a table instead of a single number (e.g. Delta with respect to to the forward rate 1M, 2M, 3M,6M..etc).This is same complexity can be seen for credit underlying as well in which the underlying is a CDS spread curve.

Bibliography

Areal, N., Rodrigues, A., and Armada, M. R. (2008). On improving the least squares monte carlo option valuation method. *Review of Derivatives Research*, 11(1-2):119.

Bachelier, L. (1900). Théorie de la spéculation. In *Annales scientifiques de l'École normale supérieure*, volume 17, pages 21–86.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Boyle, P., Broadie, M., and Glasserman, P. (1997). Monte carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21(8-9):1267–1321.

Bradtke, S. J. and Barto, A. G. (1996). Linear least-squares algorithms for temporal difference learning. *Machine Learning*, 22(1-3):33–57.

Brown, R. (1827). *Prodromus florae Novae Hollandiae et Insulae Van-Diemen: exhibens characteres plantarum*, volume 3. Sumtibus L. Schrag.

Carriere, J. F. (1996). Valuation of the early-exercise price for options using simulations and non-parametric regression. *Insurance: mathematics and Economics*, 19(1):19–30.

Clément, E., Lamberton, D., and Protter, P. (2002). An analysis of a least squares regression method for american option pricing. *Finance and Stochastics*, 6(4):449–471.

Deng, G., Mallett, J., and McCann, C. (2016). Modeling autocallable structured products. In *Derivatives and Hedge Funds*, pages 323–344. Springer.

Derman, E. and Kani, I. (1994). Riding on a smile. *Risk*, 7(2):32–39.

Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1):18–20.

Egloff, D. (2005). Monte carlo algorithms for optimal stopping and statistical learning. *The Annals of Applied Probability*, 15(2):1396–1432.

Egloff, D., Kohler, M., and Todorovic, N. (2007). A dynamic look-ahead monte carlo algorithm for pricing bermudan options. *The Annals of Applied Probability*, 17(4):1138–1171.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der Physik*, 4.

Enke, D. and Amornwattana, S. (2008). A hybrid derivative trading system based on volatility and return forecasting. *The Engineering Economist*, 53:259–292.

Ernst, D., Geurts, P., and Wehenkel, L. (2005). Tree-based batch mode reinforcement learning. *Journal of Machine Learning Research*, 6(Apr):503–556.

Feng, J., Forde, M., and Fouque, J.-P. (2010). Short-maturity asymptotics for a fast mean-reverting heston stochastic volatility model. *SIAM Journal on Financial Mathematics*, 1(1):126–141.

Forde, M., Jacquier, A., and Mijatović, A. (2010). Asymptotic formulae for implied volatility in the heston model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 466(2124):3593–3620.

Fries, C. P. and Joshi, M. S. (2008). *Conditional analytic Monte-Carlo pricing scheme of auto-callable products*. Centre for Actuarial Studies, Department of Economics, University of Melbourne.

Gamba, A. (2003). Real options valuation: A monte carlo approach. *Faculty of Management, University of Calgary WP*, (2002/3).

Gil, A., Segura, J., and Temme, N. (2007). *Numerical Methods for Special Functions*.

Glasserman, P. and Staum, J. (2001). Conditioning on one-step survival for barrier option simulations. *Operations Research*, 49(6):923–937.

Grzelak, L. A. and Oosterlee, C. W. (2011). On the heston model with stochastic interest rates. *SIAM Journal on Financial Mathematics*, 2(1):255–286.

Hachiya, H. and Sugiyama, M. (2007). Robot control by least-squares policy iteration with geodesic gaussian kernels. , 21:1–4.

Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (2002). Managing smile risk. *The Best of Wilmott*, 1:249–296.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343.

Hirsa, A. and Neftci, S. N. (2013). *An introduction to the mathematics of financial derivatives.* Academic press.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300.

Hull, J. C. (2003). *Options futures and other derivatives.* Pearson Education India.

Janzon, K. (2018). Monte carlo path simulation and the multilevel monte carlo method.

Kohler, M., Krzyżak, A., and Todorovic, N. (2010). Pricing of high-dimensional american options by neural networks. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 20(3):383–410.

Lagoudakis, M. G. and Parr, R. (2002). Model-free least-squares policy iteration. In *Advances in Neural Information Processing Systems*, pages 1547–1554.

Lemons, D. S. and Langevin, P. (2002). *An introduction to stochastic processes in physics.* JHU Press.

Li, Y., Szepesvari, C., and Schuurmans, D. (2009). Learning exercise policies for american options. In *Artificial Intelligence and Statistics*, pages 352–359.

Longstaff, F. A. and Schwartz, E. S. (2001). Valuing american options by simulation: a simple least-squares approach. *The Review of Financial Studies*, 14(1):113–147.

Marco, A. (2001). *Quantitative Analysis In Financial Markets: Collected Papers Of The New York University Mathematical Finance Seminar (Vol Ii)*, volume 2. World scientific.

Merton, R. C. (1989). On the application of the continuous-time theory of finance to financial intermediation and insurance. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 14(3):225–261.

Metropolis, N. and Ulam, S. (1949). The monte carlo method. *Journal of the American statistical association*, 44(247):335–341.

Moreno, M. and Navas, J. F. (2003). On the robustness of least-squares monte carlo (lsm) for pricing american derivatives. *Review of Derivatives Research*, 6(2):107–128.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Swishchuk, A. V. and Vadori, N. (2012). Delayed heston model: Improvement of the volatility surface fitting. *Available at SSRN 2155820*.

Tian, Y., Zhu, Z., Lee, G., Klebaner, F., and Hamza, K. (2015). Calibrating and pricing with a stochastic-local volatility model. *The Journal of Derivatives*, 22(3):21–39.

Tilley, J. A. (1993). Valuing american options in a path simulation model. In *Transactions of the Society of Actuaries*. Citeseer.

Watkins, C. J. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3-4):279–292.

Zhao, J. (2018). American option valuation methods. *International Journal of Economics and Finance*, 10:1.

Zvan, R., Vetzal, K. R., and Forsyth, P. A. (2000). Pde methods for pricing barrier options. *Journal of Economic Dynamics and Control*, 24(11-12):1563–1590.